

15/894

Delaval, Jan

From: Liu, Samuel
Sent: Wednesday, April 27, 2005 5:41 PM
To: Delaval, Jan
Subject: 09627600

Hi, Jan,

Please conduct search for the amino acid sequences of SEQ ID NO:135 against commercial protein and interference databases for application 09627600 with at least 50 search results, as timely concerned:

Very best,

Samuel Liu
AU 1653, REM 3C84
571-272-0949

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 06:03:35 ; Search time 39 Seconds
(without alignments)
17.270 Million cell updates/sec

Title: US-09-627-600D-135
Perfect score: 35
Sequence: 1 G GKARRL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 65 summaries

Database : PIR_79:*
1: pirl:*
2: pirl2:*
3: pirl3:*
4: pirl4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	33	94.3	192	2	A64098	molybdopterin-guan
2	33	94.3	194	2	B86073	hypothetical prote
3	33	94.3	194	2	S40803	molybdopterin-guan
4	33	94.3	194	2	D91226	Molybdopterin-guan
5	33	94.3	194	2	AE0951	molybdopterin-guan
6	33	94.3	357	2	S31828	band 3 anion trans
7	33	94.3	1227	2	A33638	erythrocyte anion
8	33	94.3	1232	2	I38496	anion exchanger 3
9	32	91.4	67	2	AI3268	hypothetical prote
10	32	91.4	286	2	B96615	probable carbonic
11	32	91.4	1030	2	A42497	anion exchanger 3,
12	32	91.4	1227	2	B34911	band 3-related pro
13	31	88.6	140	1	HSURBS	histone H2B.1, spe
14	31	88.6	183	2	S62700	photoassimilate-re
15	31	88.6	258	2	T35444	ABC transporter AT
16	31	88.6	354	2	E87635	efflux system prot
17	31	88.6	387	2	D88479	protein F47D12.1 l
18	31	88.6	445	2	JQ2264	canavalin - jack b
19	31	88.6	445	2	S00281	transport ATP-bind
20	31	88.6	557	2	AH3604	canavalin - sword
21	31	88.6	608	2	T34391	hypothetical prote
22	31	88.6	610	2	T19333	hypothetical prote
23	31	88.6	634	2	D70331	leucine-tRNA ligas
24	31	88.6	1096	2	C87263	hypothetical prote
25	30	85.7	145	2	T36091	probable maturase-
26	30	85.7	265	2	G95058	integrase/recombin
27	30	85.7	265	2	G97927	integrase/recombin
28	30	85.7	268	2	F87076	probable reductase
29	30	85.7	347	2	S75717	molybdenum cofacto

30	30	85.7	592	2	B48315	lamin B2 - mouse
31	30	85.7	651	2	F64417	probable ATP-depen
32	29	82.9	143	2	A96819	hypothetical prote
33	29	82.9	295	2	G70907	hypothetical prote
34	29	82.9	303	2	T36509	probable molybdopt
35	29	82.9	319	2	AC2503	integrase/recombin
36	29	82.9	327	2	D72257	hypothetical prote
37	29	82.9	359	2	A95277	hypothetical prote
38	29	82.9	385	2	F72567	hypothetical prote
39	29	82.9	396	2	T01049	hypothetical prote
40	29	82.9	417	2	T45857	hypothetical prote
41	29	82.9	454	2	T45297	tyrosine phenol-ly
42	29	82.9	588	2	T48009	PECTINESTERASE-lik
43	29	82.9	949	2	JC7802	urb protein - mous
44	29	82.9	980	2	G75523	probable cell divi
45	29	82.9	1230	2	T04181	hypothetical prote
46	29	82.9	1407	2	T28702	probable polyketid
47	29	82.9	3175	1	RRWVEV	genome polyprotein
48	29	82.9	5107	2	T29144	partial CDS - Caen
49	28	80.0	37	2	E97596	hypothetical prote
50	28	80.0	61	2	T21264	hypothetical prote
51	28	80.0	77	2	T26000	hypothetical prote
52	28	80.0	67	2	T16222	hypothetical prote
53	28	80.0	116	2	S70489	photosystem II pro
54	28	80.0	121	2	AH0895	probable membrane
55	28	80.0	139	2	S32510	homeobox-containin
56	28	80.0	169	2	T02081	ABA- and ripening-
57	28	80.0	190	2	C72666	probable mitochond
58	28	80.0	191	2	E75132	molybdopterin-guan
59	28	80.0	192	2	G71089	hypothetical prote
60	28	80.0	199	2	S41316	coat protein - cuc
61	28	80.0	249	2	AF0735	probable bacteriop
62	28	80.0	250	2	AE0621	probable prophage
63	28	80.0	281	2	G97312	site-specific DNA
64	28	80.0	288	2	H84757	hypothetical prote
65	28	80.0	317	2	T29476	hypothetical prote

ALIGNMENTS

RESULT 1
A64098 molybdopterin-guanine dinucleotide biosynthesis protein A homolog - Haemophilus influenzae
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 16-Aug-2004
C;Accession: A64098
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, .
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: A64098
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-192 <TIGR>
A;Cross-references: UNIPROT:P44899; GB:U32766; GB:L42023; NID:g1573854; PIDN:AAC22501.1;
C;Superfamily: Molybdopterin guanine dinucleotide biosynthesis protein Moba

Query Match 94.3%; Score 33; DB 2; Length 192;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 G GKARRL 7
|||||:
Db 12 G GKARRM 18

RESULT 2
B86073 hypothetical protein mobA [imported] - Escherichia coli (strain O157:H7, substrain EDL93

C/Species: Escherichia coli
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Aug-2004
C/Accession: B86073
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: A85480; MUID:21074935; PMID:11206551
A/Accession: B86073
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-194 <STO>
A/Cross-references: GB:AE005174; NID:g12518733; PIDN:AAG59046.1; GSPDB:GN00145; UWGP:Z53
A/Experimental source: strain O157:H7, substrain EDL933
C/Genetics:
A/Gene: mobA
C/Superfamily: Molybdopterin guanine dinucleotide biosynthesis protein MobA

Query Match 94.3%; Score 33; DB 2; Length 194;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
|||||:
Db 14 GGKARRM 20

RESULT 3
S40803
molybdopterin-guanine dinucleotide biosynthesis protein A - Escherichia coli (strain K-1
N/Alternate names: mob protein
C/Species: Escherichia coli
C/Date: 06-Oct-1994 #sequence_revision 26-May-1995 #text_change 12-Jul-2004
C/Accession: S40803; E65190
R/Plunkett III, G.; Burland, V.; Daniels, D.L.; Blattner, F.R.
Nucleic Acids Res. 21, 3391-3398, 1993
A/Title: Analysis of the Escherichia coli genome. III. DNA sequence of the region from 8
A/Reference number: S40802; MUID:93347969; PMID:8346018
A/Accession: S40803
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-194 <PLU>
A/Cross-references: UNIPROT:P32173; EMBL:L19201; NID:g304961; PIDN:AAB02992.1; PID:g3049
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1993
R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: B65190
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-194 <BLAT>
A/Cross-references: GB:AE000461; GB:U00096; NID:g2367318; PIDN:AAC76855.1; PID:g1790288;
A/Experimental source: strain K-12, substrain MG1655
C/Genetics:
A/Gene: mobA
A/Start codon: GTG
C/Superfamily: molybdenum cofactor biosynthesis protein, MobA type

Query Match 94.3%; Score 33; DB 2; Length 194;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
|||||:
Db 14 GGKARRM 20

RESULT 4
D91226
Molybdopterin-guanine dinucleotide biosynthesis protein A [similarity] - Escherichia col
C/Species: Escherichia coli

C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 16-Aug-2004
C/Accession: D91226
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.,
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: D91226
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-194 <HAY>
A/Cross-references: UNIPROT:P58221; GB:BA000007; PIDN:BA838203.1; PID:g13364256; GSPDB:GT
A/Experimental source: strain O157:H7, substrain RIMD 0509952
C/Genetics:
A/Gene: ECs4780
C/Superfamily: Molybdopterin guanine dinucleotide biosynthesis protein MobA

Query Match 94.3%; Score 33; DB 2; Length 194;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
|||||:
Db 14 GGKARRM 20

RESULT 5
AE0951
molybdopterin-guanine dinucleotide biosynthesis protein A [imported] - Salmonella enteric
C/Species: Salmonella enterica subsp. enterica serovar Typhi
A/Note: this species has also been called Salmonella typhi
C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 16-Aug-2004
C/Accession: AE0951
R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A/Reference number: AB0502; MUID:21534947; PMID:11677608
A/Accession: AE0951
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-194 <PAR>
A/Cross-references: GB:AL513382; PIDN:CAD03105.1; PID:g16504742; GSPDB:GN00176
C/Genetics:
A/Gene: STY3886
C/Superfamily: Molybdopterin guanine dinucleotide biosynthesis protein MobA

Query Match 94.3%; Score 33; DB 2; Length 194;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
|||||:
Db 14 GGKARRM 20

RESULT 6
S31828
band 3 anion transport protein - human (fragment)
N/Alternate names: anion exchange protein 3
C/Species: Homo sapiens (man)
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C/Accession: S31828
R/Gehrig, H.; Appelhans, H.
submitted to the EMBL Data Library, January 1993
A/Reference number: S31828
A/Accession: S31828
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-357 <GEH>
A/Cross-references: UNIPROT:Q13717; EMBL:X70797; NID:g28427; PIDN:CAA50067.1; PID:g93988

C;Superfamily: band 3 anion transport protein

Query Match 94.3%; Score 33; DB 2; Length 357;
Best Local Similarity 85.7%; Pred. No. 26;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARRL 7
|||||:
Db 138 GKGARRI 144

RESULT 7

A33638

erythrocyte anion exchanger homolog AE3 - mouse

C;Species: Mus musculus (house mouse)

C;Date: 02-Feb-1990 #sequence_revision 02-Feb-1990 #text_change 09-Jul-2004

C;Accession: A33638

R;Kopito, R.R.; Lee, B.S.; Simmons, D.M.; Lindsey, A.E.; Morgans, C.W.; Schneider, K.
Cell 59, 927-937, 1989

A;Title: Regulation of intracellular pH by a neuronal homolog of the erythrocyte anion e
A;Reference number: A33638; MUID:90075236; PMID:2686841

A;Accession: A33638

A;Status: Preliminary

A;Molecule type: mRNA

A;Residues: 1-1227 <KOP>

A;Cross-references: UNIPROT:P16283; GB:M28383; NID:g191735; PIDN:AAA37184.1; PID:g309095
C;Superfamily: band 3 anion transport protein

Query Match 94.3%; Score 33; DB 2; Length 1227;
Best Local Similarity 85.7%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARRL 7
|||||:
Db 915 GKGARRI 921

RESULT 8

I38496

anion exchanger 3 brain isoform - human

C;Species: Homo sapiens (man)

C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004

C;Accession: I38496

R;Yamoukakos, D.; Stuart-Tilley, A.; Fernandez, H.; Fey, P.; Duyk, G.; Alper, S.
Circ. Res. 75, 603-614, 1994

A;Title: Molecular cloning, expression, and chromosomal localization of two isoforms of
A;Reference number: I38496; MUID:95008042; PMID:79233606

A;Accession: I38496

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1232 <RES>

A;Cross-references: UNIPROT:P48751; EMBL:U05596; NID:g476221; PIDN:AAA50748.1; PID:g4762
C;Superfamily: band 3 anion transport protein

Query Match 94.3%; Score 33; DB 2; Length 1232;
Best Local Similarity 85.7%; Pred. No. 77;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARRL 7
|||||:
Db 920 GKGARRI 926

RESULT 9

A13268

hypothetical protein BME10134 [imported] - Brucella melitensis (strain 16M)

C;Species: Brucella melitensis

C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C;Accession: A13268

R;DelVecchio, V.G.; Kaprat, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens

A;Reference number: AD3252; PMID:11756688

A;Accession: A13268

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-67 <KUR>

A;Cross-references: UNIPROT:Q8YJ9; UNIPROT:Q8FYF1; GB:AE008917; PIDN:AAL51316.1; PID:g1
A;Experimental source: strain 16M

C;Genetics:

A;Gene: BME10134

A;Map position: 1

Query Match 91.4%; Score 32; DB 2; Length 67;
Best Local Similarity 85.7%; Pred. No. 9.8;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARRL 7
|||||:
Db 12 GKGARRL 18

RESULT 10

B96615

probable carbonic anhydrase T18124.9 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: B96615

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: B96615

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-286 <STO>

A;Cross-references: UNIPROT:Q9C6R2; GB:AE005173; NID:g11038509; PIDN:AAG27786.1; GSPDB:G

C;Genetics:

A;Gene: T18124.9

A;Map position: 1

Query Match 91.4%; Score 32; DB 2; Length 286;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARRL 7
|||||:
Db 6 GGRARRL 12

RESULT 11

A42497

anion exchanger 3, cardiac splice form - rat

N;Alternate names: AE3; chloride/bicarbonate exchanger 3

C;Species: Rattus norvegicus (Norway rat)

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C;Accession: A42497; B42497

R;Lim, S.C.; Kudrycki, K.E.; Shull, G.E.

J. Biol. Chem. 267, 7927-7935, 1992

A;Title: The predicted translation product of a cardiac AE3 mRNA contains an N terminus
gene, and identification of an alternative transcription initiation site.

A;Reference number: A42497; MUID:92218461; PMID:1560021

A;Accession: A42497

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-1030 <LIN1>

A;Cross-references: UNIPROT:Q63694; GB:M87060; NID:g202770

A;Note: this translation is not annotated in GenBank entry RA7AE3A, release 111.0; this

A/Note: sequence extracted from NCBI backbone (NCBIN:96971, NCBIPI:96975)

A/Accession: B42497

A/Molecule type: DNA

A/Residues: 1-73,'E' <LIN2>

A/Cross-references: GB:M87060; NID:g202770; PIDN:AAA40692.1; PID:g202771

A/Experimental source: tissue type spleen

C/Superfamily: band 3 anion transport protein

C/Keywords: alternative splicing; cardiac muscle; heart; transmembrane protein

Query Match

91.4%; Score 32; DB 2; Length 1030;

Best Local Similarity 85.7%; Pred. No. 1.1e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7

|||||:

Db 718 GGKARRV 724

RESULT 12

B34911 band 3-related protein 3 - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 09-Jul-2004

C/Accession: B34911

R/Kudrycki, K.E.; Newman, P.R.; Shull, G.E.

J. Biol. Chem. 265, 462-471, 1990

A/Title: cDNA cloning and tissue distribution of mRNAs for two proteins that are related

A/Reference number: A34911; MUID:90094439; PMID:2294114

A/Accession: B34911

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-1227 <KUD>

A/Cross-references: UNIPROT:P23348; GB:J05167; NID:g203088; PIDN:AAA40798.1; PID:g203089

C/Superfamily: band 3 anion transport protein

C/Keywords: transmembrane protein

Query Match

91.4%; Score 32; DB 2; Length 1227;

Best Local Similarity 85.7%; Pred. No. 1.2e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7

|||||:

Db 915 GGKARRV 921

RESULT 13

HSURBS

histone H2B.1, sperm - sea urchin (Strongylocentrotus purpuratus)

C/Species: Strongylocentrotus purpuratus (purple urchin)

C/Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004

C/Accession: S07376

R/Lai, Z.; Lieber, T.; Childs, G.

Nucleic Acids Res. 14, 9218, 1986

A/Title: The nucleotide sequence of the gene encoding the sperm specific histone subtype

A/Reference number: S07376; MUID:87066789; PMID:3786151

A/Accession: S07376

A/Molecule type: DNA

A/Residues: 1-140 <LAI>

A/Cross-references: UNIPROT:P06145; EMBL:X04681; NID:g10232; PIDN:CAA28385.1; PID:g10233

C/Superfamily: histone H2B

C/Keywords: chromosomal protein; DNA binding; nucleosome core; sperm

F/2-140/Product: histone H2B.1, sperm #status predicted <MAT>

Query Match

88.6%; Score 31; DB 1; Length 140;

Best Local Similarity 100.0%; Pred. No. 30;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARR 6

|||||:

Db 31 GGKARR 36

RESULT 14

S62700

photoassimilate-responsive protein PAR-1c precursor - common tobacco

C/Species: Nicotiana tabacum (common tobacco)

C/Date: 14-Sep-1996 #sequence_revision 07-Feb-1997 #text_change 09-Jul-2004

C/Accession: S62700; S57470

R/Herbers, K.; Moenke, G.; Badur, R.; Sonnwald, U.

Plant Mol. Biol. 29, 1027-1038, 1995

A/Title: A simplified procedure for the subtractive cDNA cloning of photoassimilate-resp

A/Reference number: S62698; MUID:96145513; PMID:8555446

A/Accession: S62700

A/Status: nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 1-183 <HER>

A/Cross-references: UNIPROT:Q43589; EMBL:X83852; NID:g871488; PIDN:CAA58732.1; PID:g87148

C/Superfamily: tobacco photoassimilate-responsive protein

F/1-28/Domain: signal sequence #status predicted <SIG>

F/29-183/Product: photoassimilate-responsive protein PAR-1c #status predicted <MAT>

Query Match

88.6%; Score 31; DB 2; Length 183;

Best Local Similarity 100.0%; Pred. No. 38;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARR 6

|||||:

Db 146 GGKARR 151

RESULT 15

T35444

ABC transporter ATP-binding protein - Streptomyces coelicolor

C/Species: Streptomyces coelicolor

C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004

C/Accession: T35444

R/Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

Submitted to the EMBL Data Library, July 1998

A/Reference number: Z21578

A/Accession: T35444

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-258 <MUR>

A/Cross-references: UNIPROT:O86745; EMBL:AL031035; PIDN:CAA19898.1; GSPDB:GN00070; SCOED

A/Experimental source: strain A3(2)

C/Genetics:

A/Gene: SCOEDB:SC6A9.13

C/Superfamily: inner membrane protein malK; ATP-binding cassette homology

Query Match

88.6%; Score 31; DB 2; Length 258;

Best Local Similarity 100.0%; Pred. No. 52;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARR 6

|||||:

Db 137 GGKARR 142

RESULT 16

E87635

efflux system protein, probable [imported] - Caulobacter crescentus

C/Species: Caulobacter crescentus

C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C/Accession: E87635

B./Nieman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

R.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A/Title: Complete Genome Sequence of Caulobacter crescentus.

A/Reference number: A87249; MUID:21173698; PMID:11259647

A/Accession: E87635

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-354 <STO>

A/Cross-references: UNIPROT:Q9A3T3; GB:AE005673; NID:g13424779; PIDN:AAK25081.1; GSPDB:G

C/Genetics:

A:Gene: CC3119

Query Match 88.6%; Score 31; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARR 6

Db 301 GGKARR 306

RESULT 17

D88479

protein F47D12.1 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: D88479

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog

A:Reference number: A75000; MUID:99069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: D88479

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-387 <STO>

A:Cross-references: UNIPROT:Q8MQ41; GB:chr_III; PIDN:AAA64313.1; PID:g722353; GSPDB:GN00

C:Genetics:

A:Gene: F47D12.1

A:Map position: 3

Query Match 88.6%; Score 31; DB 2; Length 387;
Best Local Similarity 85.7%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGKARRL 7

Db 215 GGKARRL 221

RESULT 18

JQ2264

canavalin - jack bean

C:Species: Canavalia ensiformis (jack bean)

C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004

C:Accession: JQ2264; S19137

R:Ng, J.D.; Ko, T.P.; McPherson, A.

Plant Physiol. 101, 713-728, 1993

A:Title: Cloning, expression, and crystallization of jack bean (Canavalia ensiformis) ca

A:Reference number: JQ2264; MUID:94143475; PMID:8310055

A:Accession: JQ2264

A:Molecule type: mRNA

A:Residues: 1-445 <NGJ>

A:Cross-references: UNIPROT:P50477; GB:X59467; GB:S76871; NID:g17976; PIDN:CAA42075.1; F

A:Experimental source: seed

R:Ng, J.D.; Stinchcombe, T.; Ko, T.P.; Alexander, E.; McPherson, A.

Plant Mol. Biol. 18, 147-149, 1992

A:Title: PCR cloning of the full-length cDNA for the seed protein canavalin from the jac

A:Reference number: S19137; MUID:92119225; PMID:1731967

A:Accession: S19137

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-445 <NGJ2>

A:Cross-references: EMBL:X59467; NID:g17976; PIDN:CAA42075.1; PID:g17977

C:Superfamily: glycinin

C:Keywords: storage protein

Query Match 88.6%; Score 31; DB 2; Length 445;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARR 6

Db 429 GGKARR 434

RESULT 19

S00281

canavalin - sword bean

C:Species: Canavalia gladiata (sword bean)

C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004

C:Accession: S00281; S04598

R:Yamauchi, D.; Nakamura, K.; Aeahi, T.; Minamikawa, T.

Eur. J. Biochem. 170, 515-520, 1988

A:Title: cDNAs for canavalin and concanavalin A from Canavalia gladiata seeds. Nucleoti

s.

A:Reference number: S00281; MUID:88111636; PMID:3338449

A:Accession: S00281

A:Molecule type: mRNA

A:Residues: 1-445 <YAM>

A:Cross-references: UNIPROT:P10562; EMBL:X06733; NID:g18003; PIDN:CAA29910.1; PID:g18004

R:Takei, Y.; Yamauchi, D.; Minamikawa, T.

Nucleic Acids Res. 17, 4381, 1989

A:Title: Nucleotide sequence of the canavalin gene from Canavalia gladiata seeds.

A:Reference number: S04598; MUID:89296493; PMID:2740227

A:Accession: S04598

A:Molecule type: DNA

A:Residues: 1-160,'K',162-445 <TAK>

A:Cross-references: EMBL:X15076; NID:g18006; PIDN:CAA33172.1; PID:g18007

C:Genetics:

A:Introns: 115/1; 173/3; 200/3; 290/3; 381/1

C:Superfamily: glycinin

C:Keywords: seed

Query Match 88.6%; Score 31; DB 2; Length 445;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARR 6

Db 429 GGKARR 434

RESULT 20

AH3604

transport ATP-binding protein cydC BMEI10761 [imported] - Brucella melitensis (strain 16

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C:Accession: AH3604

R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens

A:Reference number: AD3252; PMID:11756688

A:Accession: AH3604

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-557 <KUR>

A:Cross-references: UNIPROT:Q8YBX5; GB:AE008918; PIDN:AAL54003.1; PID:g17984953; GSPDB:G

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI10761

A:Map position: II

C:Superfamily: Mycobacterium tuberculosis probable ABC transporter cydC; ATP-binding cas

Query Match 88.6%; Score 31; DB 2; Length 557;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7

Db 465 GGQARRL 471

RESULT 21

T34391
hypothetical protein T26A5.1 - Caenorhabditis elegans
A/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C/Accession: T34391
R/Du, Z.
submitted to the EMBL Data Library, April 1994
A/Description: The sequence of C. elegans cosmid T26A5.
A/Reference number: Z21516
A/Accession: T34391
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-608 <DUZ>
A/Cross-references: UNIPROT:Q22802; EMBL:U00043; PIDN:AAC77504.1; GSPDB:GN00021; CESP:TT
A/Experimental source: strain Bristol N2; clone T26A5
C/Genetics:
A/Gene: CESP:T26A5.1
A/Map position: 3
A/Intons: 23/1; 96/3; 243/1; 342/2; 374/3; 403/1; 428/2; 464/3; 494/3; 534/2
C/Superfamily: fruit fly white protein; ATP-binding cassette homology

Query Match 88.6%; Score 31; DB 2; Length 608;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARRL 7
||:||||
Db 171 GGEARRL 177

RESULT 22
T19333.
hypothetical protein C16C10.12 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T19333
R/Lloyd, C.
submitted to the EMBL Data Library, November 1994
A/Reference number: Z19108
A/Accession: T19333
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-610 <WIL>
A/Cross-references: UNIPROT:Q09466; EMBL:Z46787; PIDN:CAA86750.1; GSPDB:GN00021; CESP:CL
A/Experimental source: clone C16C10
C/Genetics:
A/Gene: CESP:C16C10.12
A/Map position: 3
A/Intons: 25/3; 41/2; 76/2; 138/3; 266/3; 350/2; 441/2; 502/3; 542/2; 564/3
C/Superfamily: fruit fly white protein; ATP-binding cassette homology

Query Match 88.6%; Score 31; DB 2; Length 610;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARRL 7
||:||||
Db 182 GGEARRL 188

RESULT 23
D70331
leucine-tRNA ligase (EC 6.1.1.4) alpha chain - Aquifex aeolicus
C/Species: Aquifex aeolicus
C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C/Accession: D70331
R/Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O
V.
Nature 392, 353-358, 1998
A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A/Reference number: A70300; MUID:98196666; PMID:9537320
A/Accession: D70331
A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA
A/Residues: 1-634 <AQF>
A/Cross-references: UNIPROT:O66680; GB:AE000685; NID:g2983026; PIDN:AAC06643.1; PID:g298
A/Experimental source: strain VFS
C/Genetics:
A/Gene: leuS
C/Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 88.6%; Score 31; DB 2; Length 634;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARR 6
|||||
Db 508 GKGARR 513

RESULT 24
C87263
hypothetical protein CC0116 [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: C87263
R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: AB7249; MUID:21173698; PMID:11259647
A/Accession: C87263
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1096 <STO>
A/Cross-references: UNIPROT:Q9ABV4; GB:AE005673; NID:g13421223; PIDN:AAK22103.1; GSPDB:G
C/Genetics:
A/Gene: CC0116

Query Match 88.6%; Score 31; DB 2; Length 1096;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARRL 7
||:||||
Db 396 GGEARRL 402

RESULT 25
T36091
probable maturase-related protein - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T36091
R/Saunders, D.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, April 1999
A/Reference number: Z21596
A/Accession: T36091
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-145 <SAU>
A/Cross-references: UNIPROT:Q9X864; EMBL:AL049661; PIDN:CAB41207.1; GSPDB:GN00070; SCOE
A/Experimental source: strain A3(2)
C/Genetics:
A/Gene: SCOEDB:SCE134.09

Query Match 85.7%; Score 30; DB 2; Length 145;
Best Local Similarity 85.7%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GKGARRL 7
||| |||
Db 86 GKGARRL 92


```
RESULT 26
G95058
Integrase/recombinase, phage integrase family [imported] - Streptococcus pneumoniae (str
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C/Accession: G95058
R/Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfele,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A/Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A/Reference number: A95000; MUID:21357209; PMID:11463916
A/Accession: G95058
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-265 <KUR>
A/Cross-references: UNIPROT:Q97S82; GB:AE005672; PIDN:AAK74664.1; PID:G14971979; GSPDB:G
C/Genetics:
A/Experimental source: strain TIGR4
A/Gene: SP0506

Query Match      85.7%; Score 30; DB 2; Length 265;
Best Local Similarity 85.7%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GKGARRL 7
Db 147 GKGIRRL 153

RESULT 27
G97927
Integrase/recombinase [imported] - Streptococcus pneumoniae (strain R6)
C/Species: Streptococcus pneumoniae
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C/Accession: G97927
R/Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A/Reference number: A97872; MUID:21429245; PMID:11544234
A/Accession: G97927
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-265 <KUR>
A/Cross-references: UNIPROT:Q8DQX3; GB:AE007317; PIDN:AAK99251.1; PID:G15458015; GSPDB:G
C/Genetics:
A/Gene: xerD

Query Match      85.7%; Score 30; DB 2; Length 265;
Best Local Similarity 85.7%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GKGARRL 7
Db 147 GKGIRRL 153

RESULT 28
F87076
Probable reductase [imported] - Mycobacterium leprae
C/Species: Mycobacterium leprae
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: F87076
R/Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A/Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A/Title: Massive gene decay in the leprosy bacillus.
```

```
A/Reference number: A86909; MUID:21128732; PMID:11234002
A/Accession: F87076
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-268 <STO>
A/Cross-references: UNIPROT:O05684; GB:AL450380; NID:G13093249; PIDN:CAC31721.1; GSPDB:G
C/Genetics:
A/Gene: ML1340

Query Match      85.7%; Score 30; DB 2; Length 268;
Best Local Similarity 71.4%; Pred. No. 87;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKGARRL 7
Db 231 GGRARRI 237

RESULT 29
S75717
molybdenum cofactor biosynthesis protein C - Synechocystis sp. (strain PCC 6803)
N/Alternate names: protein slr0902
C/Species: Synechocystis sp.
A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 12-Jul-2004
C/Accession: S75717
R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
S.
A/Reference number: S74322; MUID:97061201; PMID:8905231
A/Accession: S75717
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-347 <KAN>
A/Cross-references: UNIPROT:Q55370; EMBL:D64003; GB:AB001339; NID:G1001200; PIDN:BA1045.
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C/Genetics:
A/Gene: moaC
C/Superfamily: bifunctional molybdenum cofactor biosynthesis protein MoaC/MobA

Query Match      85.7%; Score 30; DB 2; Length 347;
Best Local Similarity 71.4%; Pred. No. 11e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKGARRL 7
Db 169 GGSRRM 175

RESULT 30
B48315
lamin B2 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C/Accession: B48315; S08264; A41068; A56583; S21609
R/Hoeger, T.H.; Zatloukal, K.; Waizenegger, I.; Krohne, G.
Chromosoma 99, 379-390, 1990
A/Title: Characterization of a second highly conserved B-type lamin present in cells pre
A/Reference number: A48315; MUID:91106216; PMID:2102682
A/Accession: B48315
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-592 <HOE>
A/Cross-references: UNIPROT:P21619; EMBL:X54098; NID:G52866; PIDN:CAA38032.1; PID:G52867
A/Note: the figure legends for Fig. 5 (African clawed frog) and Fig. 6 (mouse) appear to
sequences in GenBank
R/Hoeger, T.H.; Zatloukal, K.; Waizenegger, I.; Krohne, G.
Chromosoma 100, 67-69, 1990
A/Reference number: A56583; MUID:91339548; PMID:2102440
A/Contents: annotation; erratum
A/Note: corrects transposition of Figs. 5 and 6
```

R;Weber, K.; Plessmann, U.; Traub, P.
FEBS Lett. 261, 361-364, 1990
A/Title: Protein chemical analysis of purified murine lamin B identifies two distinct po
A/Reference number: S08264; MUID:90184481; PMID:2311764
A/Accession: S08264
A/Status: preliminary
A/Molecule type: protein
A/Residues: 182-206;234-286;290-313,'R',315-317,'X',319;364-401;466-479,'X',481-492 <WE
R;Kasahara, K.; Chida, K.; Tsunenaga, M.; Kohno, Y.; Ikuta, T.; Kuroki, T.
J. Biol. Chem. 266, 20018-20023, 1991
A/Title: Identification of lamin B-2 as a substrate of protein kinase C in BALB/MK-2 mou
A/Reference number: A41068; MUID:92041823; PMID:1939065
A/Accession: A41068
A/Molecule type: protein
A/Residues: 165-166,'X',168,'X',170-172;254-258,'X',260-262;478-479,'X',481-483,'X',485-
C;Superfamily: cytoskeletal keratin
C/Keywords: nucleus

Query Match 85.7%; Score 30; DB 2; Length 592;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGKARRL 7
Db 412 GGKRRRL 418

RESULT 31
F64417
Probable ATP-dependent helicase - Methanococcus jannaschii
C/Species: Methanococcus jannaschii
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C/Accession: F64417
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
A/Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A/Reference number: A64300; MUID:96337999; PMID:8688087
A/Accession: F64417
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-651 <BUL>
A/Cross-references: UNIPROT:Q58352; GB:U67537; GB:L77117; NID:g2826351; PIDN:AAB98945.1;
C/Genetics:
A/Map position: FOR871495-873450
C/Keywords: ATP, nucleotide binding, P-loop
F;41-48/Region: nucleotide-binding motif A (P-loop)
F;191-196/Region: nucleotide-binding motif B
F;195-198/Region: DEAH motif

Query Match 85.7%; Score 30; DB 2; Length 651;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGKARRL 7
Db 106 GGGANRL 112

RESULT 32
A96819
hypothetical protein F9K20.1 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: A96819
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, I
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: A96819
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-143 <STO>
A/Cross-references: UNIPROT:Q9ZVB2; GB:AE005173; NID:g3834301; PIDN:AAC83017.1; GSPDB:GN
C/Genetics:
A/Gene: F9K20.1
A/Map position: 1

Query Match 82.9%; Score 29; DB 2; Length 143;
Best Local Similarity 71.4%; Pred. No. 82;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
Db 115 GGSRRV 121

RESULT 33
G70907
hypothetical protein RV0588 - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: G70907
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: G70907
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-295 <COL>
A/Cross-references: UNIPROT:O07790; GB:Z97182; GB:AL123456; NID:g3250720; PIDN:CAB09950.
A/Experimental source: strain H37RV
C/Genetics:
A/Gene: RV0588
C;Superfamily: conserved hypothetical protein H11086

Query Match 82.9%; Score 29; DB 2; Length 295;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGKARRL 7
Db 29 GGTARRL 35

RESULT 34
T36509
probable molybdopterin-guanine dinucleotide biosynthesis protein - Streptomyces coelicol
C/Species: Streptomyces coelicolor
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T36509
R;Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1999
A/Reference number: Z21608
A/Accession: T36509
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-303 <SAU>
A/Cross-references: UNIPROT:Q9XA50; EMBL:AL096822; PIDN:CAB46951.1; GSPDB:GN00070; SCOE
A/Experimental source: strain A3(2)
C/Genetics:
A/Gene: SCOEDB:SCGD3.29c

Query Match 82.9%; Score 29; DB 2; Length 303;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGKARRL 7
|||
Db 4 GGAARRL 10

RESULT 35

AC2503
Integrase/recombinase [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120alpha
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AC2503
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi, S.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AC2503
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-319 <KUR>
A;Cross-references: UNIPROT:Q8YKT8; GB:BA000020; PIDN:BAE78287.1; PID:gl17135741; GSPDB:C
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr7203
A;Genome: plasmid

Query Match 82.9%; Score 29; DB 2; Length 319;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGKARRL 7
|||
Db 192 GGGKRRV 198

RESULT 36

D72257
hypothetical protein TM1403 - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 16-Aug-2004
C;Accession: D72257
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.

Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: D72257
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-327 <ARN>
A;Cross-references: UNIPROT:Q9X1C3; GB:AE001793; GB:AE000512; NID:g4981963; PIDN:AAD3647
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1403
C;Superfamily: ATP-binding cassette homology
F;20-212/Domain: ATP-binding cassette homology <ABC>

Query Match 82.9%; Score 29; DB 2; Length 327;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGKARRL 7
|||
Db 138 GGGMARRL 144

RESULT 37

A95277
hypothetical protein Sma0226 [imported] - Sinorhizobium meliloti (strain 1021) magaplast
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: A95277

R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowe, K.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C. Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A;Reference number: A95262; MUID:21396509; PMID:11481432

A;Accession: A95277
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <KUR>
A;Cross-references: UNIPROT:Q930S4; GB:AE006469; PIDN:AAK64779.1; PID:gl14523187; GSPDB:C
A;Experimental source: strain 1021, megaplastid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; D., Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: Sma0226
A;Genome: plasmid

Query Match 82.9%; Score 29; DB 2; Length 359;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGKARRL 7
|||
Db 94 GGGRRRV 100

RESULT 38

F72567
hypothetical protein APE1824 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: F72567
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: F72567
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-385 <KAW>
A;Cross-references: UNIPROT:Q9YAX1; DDBJ:AP000062; NID:g5105244; PIDN:BAA80827.1; PID:g5
C;Genetics:
A;Gene: APE1824
C;Superfamily: Pyrococcus abyssi hypothetical protein PAB1513

Query Match 82.9%; Score 29; DB 2; Length 385;
Best Local Similarity 71.4%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGKARRL 7
|||
Db 91 GGGRRRM 97

RESULT 39

T01049
hypothetical protein YUP8H12R.32 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004

C/Accession: T01049
R/Theologis, A.; Vysotskaia, V.S.; Osborne, B.I.; Schwartz, J.R.; Federspiel, N.A.; Kwan
Oefner, P.; Davis, R.W.
submitted to the EMBL Data Library, May 1998
A/Description: Arabidopsis thaliana chromosome 1 YAC YUP8H12R sequence.
A/Reference number: Z14227
A/Accession: T01049
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-396 <THE>
A/Cross-references: UNIPROT:O64545; EMBL:AC002986; NID:g2494106; PID:g3152569; GSPDB:GNC
C/Genetics:
A/Gene: ATSP:YUP8H12R.32
A/Map position: 1

Query Match 82.9%; Score 29; DB 2; Length 396;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GKGARRL 7
 |||||
 :
Db 91 GKGCRRM 97

RESULT 40
T45857
hypothetical protein F3A4.120 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C/Accession: T45857
R/Bargues, M.; Collado, M.C.; Navarro, P.; Terol, J.; Perez-Alonso, M.; Mewes, H.W.; May
submitted to the Protein Sequence Database, December 1999
A/Reference number: Z23007
A/Accession: T45857
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-417 <BAR>
A/Cross-references: UNIPROT:Q9SN14; EMBL:AL132978
A/Experimental source: cultivar Columbia; BAC clone F3A4
C/Genetics:
A/Map position: 3
A/Introns: 211/1; 226/1; 362/3
A/Note: F3A4.120
C/Superfamily: Arabidopsis thaliana hypothetical protein F3A4.120

Query Match 82.9%; Score 29; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GKGARRL 7
 |||||
 :
Db 264 GKGARRL 269

RESULT 41
T45297
tyrosine phenol-lyase (EC 4.1.99.2) [imported] - Rhodobacter capsulatus
C/Species: Rhodobacter capsulatus
C/Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004
C/Accession: T45297
R/Cortez, N.; Carrillo, N.; Pasternak, C.; Balzer, A.; Klug, G.
J. Bacteriol. 180, 5413-5420, 1998
A/Title: Molecular cloning and expression analysis of the Rhodobacter capsulatus sodB ge
A/Reference number: Z22962; MUID:98440433; PMID:9765573
A/Accession: T45297
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-454 <COR>
A/Cross-references: UNIPROT:O30971; EMBL:AF022932; NID:g2511750; PIDN:AAC64208.1; PID:g2
A/Experimental source: strain 37b4
C/Genetics:
A/Gene: tp1
C/Superfamily: tryptophanase

C/Keywords: carbon-carbon lyase

Query Match 82.9%; Score 29; DB 2; Length 454;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GKGARRL 7
 |||||
 :
Db 111 GKGRRRI 117

RESULT 42
T48009
PECTINESTERASE-like protein - Arabidopsis thaliana
N/Alternate names: protein T17J13.130
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C/Accession: T48009
R/Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Lemcke, K.; Mayer,
submitted to the Protein Sequence Database, February 2000
A/Reference number: Z24482
A/Accession: T48009
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-588 <RIE>
A/Cross-references: UNIPROT:Q9M1Q7; EMBL:AL138651
A/Experimental source: cultivar Columbia; BAC clone T17J13
C/Genetics:
A/Map position: 3
A/Introns: 360/1
A/Note: T17J13.130
C/Superfamily: pectinesterase

Query Match 82.9%; Score 29; DB 2; Length 588;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GKGARRL 7
 |||||
 :
Db 221 GGAARRL 227

RESULT 43
JC7802
Urb protein - mouse
C/Species: Mus musculus (house mouse)
C/Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004
C/Accession: JC7802
R/Aoki, K.; Sun, Y.; Aoki, S.; Wada, K.; Wada, E.
Biochem. Biophys. Res. Commun. 290, 1282-1288, 2002
A/Title: Cloning, expression, and mapping of a gene that is upregulated in adipose tissu
A/Reference number: JC7802; PMID:11812002; MUID:21670972
A/Accession: JC7802
A/Molecule type: mRNA
A/Residues: 1-949 <AOX>
A/Cross-references: UNIPROT:Q8R2G6; DDBJ:AB075019
A/Note: Three ATG codons are present in this ORF, the first Met residue is designated "p
C/Comment: This protein, a secretory protein, whose expression is tissue-specific, plays
C/Genetics:
A/Gene: Urb
A/Map position: 16
C/Keywords: adipose tissue
F;1-23/Domain: highly hydrophobic, signal sequence #status predicted <SIG>
F;282-608/Region: lysine-rich #status predicted
F;523-524/Region: cleavage recognition site by furin (Arg-Arg) #status predicted

Query Match 82.9%; Score 29; DB 2; Length 949;
Best Local Similarity 71.4%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GKGARRL 7
 |||||
 :
Db 197 GKGVRRI 203


```
RESULT 44
G75523
probable cell division protein FtsK - Deinococcus radiodurans (strain R1)
C/Species: Deinococcus radiodurans
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: G75523
R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.U.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A/Reference number: A75250; MUID:20036896; PMID:10567266
A/Accession: G75523
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-980 <WHI>
A/Cross-references: UNIPROT:Q9RXB5; GB:AE001900; GB:AE000513; NID:g6458079; PIDN:AAF0998
A/Experimental source: strain R1
C/Genetics:
A/Gene: DR0400
A/Map position: 1

Query Match      82.9%; Score 29; DB 2; Length 980;
Best Local Similarity 71.4%; Pred. No. 4.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGAARL 7
Db      99 GGAARL 105

RESULT 45
T04181
hypothetical protein F7L13.40 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C/Accession: T04181
R/Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X
submitted to the Protein Sequence Database, March 1999
A/Reference number: Z15184
A/Accession: T04181
A/Molecule type: DNA
A/Residues: 1-1230 <BEV>
A/Cross-references: UNIPROT:Q9S2Y0; EMBL:AL049524
A/Experimental source: cultivar Columbia; BAC clone F7L13
C/Genetics:
A/Map position: 4
A/Introns: 349/3; 726/2; 759/3; 1070/3
A/Note: F7L13.40
C/Superfamily: retrovirus-related polypeptide

Query Match      82.9%; Score 29; DB 2; Length 1230;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GGAARL 7
Db      463 GGAARL 468

RESULT 46
T28702
probable polyketide synthetase [similarity] - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T28702
R/Parkhill, J.; Bentley, S.D.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, January 1999
A/Reference number: Z20512
A/Accession: T28702
A/Status: preliminary; translated from GB/EMBL/DBJ
```

```
A/Molecule type: DNA
A/Residues: 1-1407 <PAR>
A/Cross-references: UNIPROT:O69826; EMBL:AL023496; PIDN:CAA18919.1
C/Superfamily: enterobactin synthetase component F; acetate-CoA ligase homology; acyl ce
C/Keywords: carrier protein; phosphopantetheine; phosphoprotein
F;591-1049/Domain: acetate-CoA ligase homology <ACT>
F;1077-1145/Domain: acyl carrier protein homology <ACP>
F;1109/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match      82.9%; Score 29; DB 2; Length 1407;
Best Local Similarity 85.7%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GGAARL 7
Db      1388 GGAARL 1394

RESULT 47
RMWV6V
genome polypeptide - equine arteritis virus
N/Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C/Species: equine arteritis virus
A/Note: host Equus caballus (domestic horse)
C/Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C/Accession: A39925; S10158; B39925
R/Den Boon, J.A.; Snijder, E.J.; Chirnside, E.D.; De Vries, A.A.F.; Horzinek, M.C.; Spaa
J. Virol. 65, 2910-2920, 1991
A/Title: Equine arteritis virus is not a togavirus but belongs to the coronaviruslike su
A/Reference number: A39925; MUID:91237805; PMID:1851863
A/Accession: A39925
A/Molecule type: genomic RNA
A/Residues: 1-3175 <DEN>
A/Cross-references: UNIPROT:P19811; EMBL:X53459
A/Note: a -1 ribosomal frameshift occurs between the codons AAC for 1727-Asn and CUG for
R/de Vries, A.A.F.; Chirnside, E.D.; Bredenbeek, P.J.; Gravestein, L.A.; Horzinek, M.C.;
Nucleic Acids Res. 18, 3241-3247, 1990
A/Title: All subgenomic mRNAs of equine arteritis virus contain a common leader sequence
A/Reference number: S10158; MUID:90287699; PMID:2162519
A/Accession: S10158
A/Status: translation not shown
A/Molecule type: genomic RNA
A/Residues: 1-17 <VRI>
A/Cross-references: EMBL:X52277
C/Superfamily: equine arteritis virus RNA-directed RNA polymerase
C/Keywords: nucleotidyltransferase

Query Match      82.9%; Score 29; DB 1; Length 3175;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGAARL 7
Db      1848 GGAARL 1854

RESULT 48
T29144
partial CDS - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T29144
R/Pauley, A.; Gattung, S.
submitted to the EMBL Data Library, July 1996
A/Description: The sequence of C. elegans cosmid K11C4.
A/Reference number: Z20577
A/Accession: T29144
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-5107 <PAU>
A/Cross-references: UNIPROT:Q94279; EMBL:U64854; PIDN:AAB18318.1; GSPDB:GN00023; CESP:unc
A/Experimental source: strain Bristol N2; clone K11C4
C/Genetics:
```

Job time : 44 secs

A;Gene: CESP:unc-68

A;Map position: 5

A;Introns: 27/1; 64/3; 92/3; 127/1; 158/2; 1222/2; 1300/2; 1347/2; 1391/1; 1419/3; 1517/
/3; 3269/2; 3313/2; 3466/1; 3519/3; 3615/3; 3629/3; 3658/2; 3710/1; 3741/3; 3779/2; 3810
C;Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homold

Query Match 82.9%; Score 29; DB 2; Length 5107;

Best Local Similarity 71.4%; Pred. No. 1.9e+03;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7

Db 639 GGKSRKV 645

RESULT 49

E97596

hypothetical protein AGR_C_3586 [imported] - Agrobacterium tumefaciens (strain C58, Cere

C;Species: Agrobacterium tumefaciens

C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004

C;Accession: E97596

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A;Reference number: A97359; MUID:21608551; PMID:11743194

A;Accession: E97596

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-37 <KUR>

A;Cross-references: UNIPROT:Q8U568; GB:AE007869; PIDN:AAK87726.1; PID:g15157091; GSPDB:G

C;Genetics:

A;Gene: AGR_C_3586

A;Map position: circular chromosome

Query Match 80.0%; Score 28; DB 2; Length 37;

Best Local Similarity 71.4%; Pred. No. 41;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGKARRL 7

Db 7 GGRTTTRL 13

RESULT 50

T21264

hypothetical protein F22D6.13 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T21264

R;Wilkinson, J.

submitted to the EMBL Data Library, April 1996

A;Reference number: Z19397

A;Accession: T21264

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-61 <WIL>

A;Cross-references: EMBL:Z71262; PIDN:CAA95819.1; GSPDB:GN00019; CESP:F22D6.13

A;Experimental source: clone F22D6

C;Genetics:

A;Gene: CESP:F22D6.13

A;Map position: 1

Query Match 80.0%; Score 28; DB 2; Length 61;

Best Local Similarity 83.3%; Pred. No. 64;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARR 6

Db 47 GGKSRV 52

Search completed: April 28, 2005, 06:26:09

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 06:02:11 ; Search time 65 Seconds
(without alignments)
55.147 Million cell updates/sec

Title: US-09-627-600D-135
Perfect score: 35
Sequence: 1 GKGARRL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 65 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	94.3	192	1	MOBA_HAEIN
2	33	94.3	194	1	MOBA_ECOS7
3	33	94.3	194	1	MOBA_ECOL6
4	33	94.3	194	1	MOBA_ECOL1
5	33	94.3	194	1	MOBA_SALTI
6	33	94.3	194	1	MOBA_SALTY
7	33	94.3	194	2	Q83PF7
8	33	94.3	357	2	Q13717
9	33	94.3	357	2	Q99416
10	33	94.3	1030	2	Q9ERP4
11	33	94.3	1227	1	B3A3_MOUSE
12	33	94.3	1227	2	Q9ERF5
13	33	94.3	1232	1	B3A3_HUMAN
14	33	94.3	1233	1	B3A3_RABIT
15	33	94.3	1239	2	Q68EG4
16	33	94.3	1259	2	Q6YIQ9
17	32	91.4	67	2	Q8YJES
18	32	91.4	67	2	Q8FYF1
19	32	91.4	179	2	Q8GVP5
20	32	91.4	234	2	Q62L10
21	32	91.4	255	2	Q6L833
22	32	91.4	286	2	Q9C6R2
23	32	91.4	290	2	Q9C6F5
24	32	91.4	304	2	Q8TXN6
25	32	91.4	356	2	Q85IG5
26	32	91.4	416	2	Q7WMD3
27	32	91.4	420	2	Q94H89
28	32	91.4	431	2	Q45081
29	32	91.4	487	2	Q7WT75
30	32	91.4	492	2	Q8QNN6
31	32	91.4	548	2	Q9EWF4

32	32	91.4	576	2	Q65WZ7	Q65wz7 oryza sativ
33	32	91.4	701	2	Q92954	Q92954 rous sarcom
34	32	91.4	1227	1	B3A3_RAT	P23348 rattus norv
35	31	88.6	48	2	Q8U2F1	Q8u2f1 pyrococcus
36	31	88.6	59	2	Q6UKB6	Q6ukb6 burkholderi
37	31	88.6	59	2	Q8HAM5	Q8ham5 burkholderi
38	31	88.6	84	2	Q6Z463	Q6z463 oryza sativ
39	31	88.6	139	1	H2B1_STRPU	P06145 strongyloce
40	31	88.6	160	2	Q8LH66	Q8lh66 oryza sativ
41	31	88.6	165	2	Q8TY99	Q8ty99 methanopyru
42	31	88.6	173	2	Q7U8G8	Q7u8g8 synecococc
43	31	88.6	183	2	Q43589	Q43589 nicotiana t
44	31	88.6	215	2	Q9AWL3	Q9awl3 oryza sativ
45	31	88.6	246	2	Q7NKD9	Q7nkd9 gloeobacter
46	31	88.6	258	2	Q86745	Q86745 streptomyce
47	31	88.6	262	2	Q8LJN4	Q8ljn4 oryza sativ
48	31	88.6	264	2	Q8S3R5	Q8s3r5 oryza sativ
49	31	88.6	299	2	Q7QER1	Q7qer1 anopheles g
50	31	88.6	315	2	Q6NK28	Q6nk28 corynebacte
51	31	88.6	322	2	Q9FDD2	Q9fdd2 brucella ab
52	31	88.6	334	2	Q8I7J8	Q8i7j8 caenorhabdi
53	31	88.6	354	2	Q9A3T3	Q9a3t3 caulobacter
54	31	88.6	409	2	Q7XAA4	Q7xaa4 oryza sativ
55	31	88.6	437	2	Q7NYK3	Q7nyk3 chromobacte
56	31	88.6	445	1	CANA_CANEN	P50477 canavalia e
57	31	88.6	445	1	CANA_CANGI	P10562 canavalia g
58	31	88.6	549	2	Q6AH59	Q6ah59 leifsonia x
59	31	88.6	557	2	Q8YBX5	Q8ybx5 brucella me
60	31	88.6	560	2	Q8FWE6	Q8fwe6 brucella su
61	31	88.6	608	2	Q22802	Q22802 caenorhabdi
62	31	88.6	610	1	YQ5C_CAEEL	Q09466 caenorhabdi
63	31	88.6	612	2	Q7Q7G5	Q7q7g5 anopheles g
64	31	88.6	627	1	ACM2_CAEEL	Q09388 caenorhabdi
65	31	88.6	634	1	SYLA_AQUAE	O66680 aquifex aeo

ALIGNMENTS

RESULT 1	ID	MOBA_HAEIN	STANDARD;	PRT;	192 AA.
AC	P44899;				
DT	01-NOV-1995 (Rel. 32, Created)				
DT	01-NOV-1995 (Rel. 32, Last sequence update)				
DT	25-OCT-2004 (Rel. 45, Last annotation update)				
DE	Probable molybdopter-in-guanine dinucleotide biosynthesis protein A.				
GN	Name=moba; Synonyms=mob; Ordered locusNames=HI0844;				
OS	Haemophilus influenzae.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;				
OC	Pasteurellaceae; Haemophilus.				
OX	NCBI_TaxID=727;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Rd / KM20 / ATCC 51907;				
RX	MEDLINE=95350630; PubMed=7542800;				
RA	Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,				
RA	Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,				
RA	McKenney K., Sutton G.G., FitzHugh W., Fields C.A., Gocayne J.D.,				
RA	Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,				
RA	Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,				
RA	Uterback T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,				
RA	Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,				
RA	Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,				
RA	Venter J.C.;				
RT	"Whole-genome random sequencing and assembly of Haemophilus influenzae				
RT	Rd.";				
RL	Science 269:496-512(1995).				
CC	-1- FUNCTION: Links a guanosine 5'-phosphate to molybdopter-in (MPT)				
CC	forming molybdopter-in guanine dinucleotide (MGD) (By similarity).				
CC	-1- PATHWAY: Molybdenum cofactor biosynthesis.				
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).				
CC	-1- SIMILARITY: Belongs to the mobA family.				

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32766; AAC2501.1; -.
DR PIR; A64098; A64098.
DR HSSP; P32173; 1E5K.
DR TIGR; H10844; -.
DR HAMAP; MF_00316; -.
KW Complete proteome; GTP-binding; Molybdenum cofactor biosynthesis.
SQ SEQUENCE 192 AA; 21754 MW; 804AC64948CD3AFD CRC64;

Query Match          94.3%; Score 33; DB 1; Length 192;
Best Local Similarity 85.7%; Pred. No. 55;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKXARRL 7
Db 12 GKXARRM 18

RESULT 2
MOBA_ECO57
ID MOBA_ECO57 STANDARD; PRT; 194 AA.
AC P58221;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Molybdopterin-guanine dinucleotide biosynthesis protein A (Protein
DE FA).
GN Name=mobA; Synonyms=chlB, mob, narB; OrderedLocusNames=z5389, ECGs4780;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoustis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [3]
RP SEQUENCE OF 87-194 FROM N.A.
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RA Ohnishi M., Murata T., Nakayama K., Kuhara S., Hattori M.,
RA Kurokawa K., Yasunaga T., Makino K., Shinagawa H., Hayashi T.;
RT "Development of primer sets for direct sequence determination of all
RT the ribosomal operons of Escherichia coli.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Links a guanosine 5'-phosphate to molybdopterin (MPT)
CC forming molybdopterin guanine dinucleotide (MGD) (By similarity).
```

```
CC -!- COFACTOR: Requires magnesium for activity (By similarity).
CC -!- PATHWAY: Molybdenum cofactor biosynthesis.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the mobA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005616; AAC59046.1; -.
DR EMBL; AP002567; BAB38203.1; -.
DR EMBL; AB035920; BAA93557.1; -.
DR PIR; D91226; D91226.
DR HSSP; P32173; 1FR9.
DR HAMAP; MF_00316; -.
KW Complete proteome; GTP-binding; Magnesium;
KW Molybdenum cofactor biosynthesis.
SQ SEQUENCE 194 AA; 21629 MW; E6F2ABDEC7BA2205 CRC64;

Query Match          94.3%; Score 33; DB 1; Length 194;
Best Local Similarity 85.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKXARRL 7
Db 14 GKXARRM 20

RESULT 3
MOBA_ECOL6
ID MOBA_ECOL6 STANDARD; PRT; 194 AA.
AC Q8FBH7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Molybdopterin-guanine dinucleotide biosynthesis protein A (Protein
DE FA).
GN Name=mobA; OrderedLocusNames=c4801;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Raske D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -!- FUNCTION: Links a guanosine 5'-phosphate to molybdopterin (MPT)
CC forming molybdopterin guanine dinucleotide (MGD) (By similarity).
CC -!- COFACTOR: Requires magnesium for activity (By similarity).
CC -!- PATHWAY: Molybdenum cofactor biosynthesis.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the mobA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC -----
DR EMBL; AE016770; AAN83232.1; -.
DR HSSP; P32173; 1E5K.
DR HAMAP; MF_00316; -. 1.
KW Complete proteome; GTP-binding; Magnesium;
KW Molybdenum cofactor biosynthesis.
SQ SEQUENCE 194 AA; 21773 MW; 2720A8F03FD373C5 CRC64;

Query Match          94.3%; Score 33; DB 1; Length 194;
Best Local Similarity 85.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAARRL 7
   |||||:
Db 14 GGAARRM 20

RESULT 4
MOBA_ECOLI STANDARD; PRT; 194 AA.
ID MOBA_ECOLI
AC P32173; Q9LBV0;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Molybdopterin-guanine dinucleotide biosynthesis protein A (Protein
DE PA).
GN Name=moba; Synonyms=chlB, mob, narB; OrderedLocusNames=b3857;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=93347969; PubMed=8346018;
RA Plunkett G. III, Burland V.D., Daniels D.L., Blattner F.R.;
RT "Analysis of the Escherichia coli genome. III. DNA sequence of the
RT region from 87.2 to 89.2 minutes.";
RL Nucleic Acids Res. 21:3391-3398(1993).
RN [2]
RP SEQUENCE OF 1-5, AND CHARACTERIZATION.
RX MEDLINE=94291668; PubMed=8020507;
RA Palmer T., Vasishta A., Whitty P.W., Boxer D.H.;
RT "Isolation of protein FA, a product of the mob locus required for
RT molybdenum cofactor biosynthesis in Escherichia coli.";
RL Eur. J. Biochem. 222:687-692(1994).
RN [3]
RP CHARACTERIZATION.
RC STRAIN=K12;
RX MEDLINE=96004466; PubMed=7551035;
RA Iobbi-Nivol C., Palmer T., Whitty P.W., McNairn E., Boxer D.H.;
RT "The mob locus of Escherichia coli K12 required for molybdenum
RT cofactor biosynthesis is expressed at very low levels.";
RL Microbiology 141:1663-1671(1995).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS).
RX MEDLINE=20568278; PubMed=10978347; DOI=10.1074/jbc.M007406200;
RA Lake M.W., Temple C.A., Rajagopalan K.V., Schindelin H.;
RT "The crystal structure of the Escherichia coli Moba protein provides
RT insight into molybdopterin guanine dinucleotide biosynthesis.";
RL J. Biol. Chem. 275:40211-40217(2000).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.35 ANGSTROMS).
RX PubMed=11080634; DOI=10.1016/S0969-2126(00)00518-9;
RA Stevenson C.E., Sargent F., Buchanan G., Palmer T., Lawson D.M.;
RT "Crystal structure of the molybdenum cofactor biosynthesis protein
RT Moba from Escherichia coli at near-atomic resolution.";
RL Structure 8:1115-1125(2000).
CC -!- FUNCTION: Links a guanosine 5'-phosphate to molybdopterin (MPT)
CC forming molybdopterin guanine dinucleotide (MGD).
CC -!- COFACTOR: Requires magnesium for activity.
CC -!- PATHWAY: Molybdenum cofactor biosynthesis.
CC -!- SUBUNIT: Monomer.

```

```

CC -----
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the mobA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L19201; AAB02992.1; -.
DR EMBL; U00096; AAC76855.1; -.
DR PIR; S40803; S40803.
DR PDB; 1E5K; X-ray; A=1-194.
DR PDB; 1FR9; X-ray; A=1-194.
DR PDB; 1FRW; X-ray; A=1-194.
DR PDB; 1H4C; X-ray; A=1-194.
DR PDB; 1H4D; X-ray; A=1-194.
DR PDB; 1H4E; X-ray; A=1-194.
DR PDB; 1HJ; X-ray; A=1-194.
DR PDB; 1HJL; X-ray; A=1-194.
DR Ecobase; EB1776; -.
DR EcoGene; EG11829; mobA.
DR HAMAP; MF_00316; -. 1.
KW 3D-structure; Complete proteome; Direct protein sequencing;
KW GTP-binding; Magnesium; Molybdenum cofactor biosynthesis.
RN [1]
RP SEQUENCE
RC STRAIN=K12
RX MEDLINE=93347969; PubMed=8346018;
RA Plunkett G. III, Burland V.D., Daniels D.L., Blattner F.R.;
RT "Analysis of the Escherichia coli genome. III. DNA sequence of the
RT region from 87.2 to 89.2 minutes.";
RL Nucleic Acids Res. 21:3391-3398(1993).
RN [2]
RP SEQUENCE OF 1-5, AND CHARACTERIZATION.
RX MEDLINE=94291668; PubMed=8020507;
RA Palmer T., Vasishta A., Whitty P.W., Boxer D.H.;
RT "Isolation of protein FA, a product of the mob locus required for
RT molybdenum cofactor biosynthesis in Escherichia coli.";
RL Eur. J. Biochem. 222:687-692(1994).
RN [3]
RP CHARACTERIZATION.
RC STRAIN=K12;
RX MEDLINE=96004466; PubMed=7551035;
RA Iobbi-Nivol C., Palmer T., Whitty P.W., McNairn E., Boxer D.H.;
RT "The mob locus of Escherichia coli K12 required for molybdenum
RT cofactor biosynthesis is expressed at very low levels.";
RL Microbiology 141:1663-1671(1995).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS).
RX MEDLINE=20568278; PubMed=10978347; DOI=10.1074/jbc.M007406200;
RA Lake M.W., Temple C.A., Rajagopalan K.V., Schindelin H.;
RT "The crystal structure of the Escherichia coli Moba protein provides
RT insight into molybdopterin guanine dinucleotide biosynthesis.";
RL J. Biol. Chem. 275:40211-40217(2000).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.35 ANGSTROMS).
RX PubMed=11080634; DOI=10.1016/S0969-2126(00)00518-9;
RA Stevenson C.E., Sargent F., Buchanan G., Palmer T., Lawson D.M.;
RT "Crystal structure of the molybdenum cofactor biosynthesis protein
RT Moba from Escherichia coli at near-atomic resolution.";
RL Structure 8:1115-1125(2000).
CC -!- FUNCTION: Links a guanosine 5'-phosphate to molybdopterin (MPT)
CC forming molybdopterin guanine dinucleotide (MGD).
CC -!- COFACTOR: Requires magnesium for activity.
CC -!- PATHWAY: Molybdenum cofactor biosynthesis.
CC -!- SUBUNIT: Monomer.

```



```
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Molybdopterin-guanine dinucleotide biosynthesis protein A (Protein
DE FA).
GN Name=moba; OrderedLocusNames=STY3886, t3626;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RX DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
CC -!- FUNCTION: Links a guanosine 5'-phosphate to molydopterin (MPT)
CC forming molybdopterin guanine dinucleotide (MGD) (By similarity).
CC -!- COFACTOR: Requires magnesium for activity (By similarity).
CC -!- PATHWAY: Molybdenum cofactor biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the mobA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL627280; CAD03105.1; -.
DR EMBL; AE016846; AAO71127.1; -.
DR HSSP; P32173; 1FR9.
DR HAMAP; MF_00316; -.
KW Complete proteome; GTP-binding; Magnesium;
KW Molybdenum cofactor biosynthesis.
SQ SEQUENCE 194 AA; 21256 MW; 81961E9C2DE4FB70 CRC64;

Query Match 94.3%; Score 33; DB 1; Length 194;
Best Local Similarity 85.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
Db 14 GGKARRM 20

RESULT 6
MOBA_SALTY STANDARD; PRT; 194 AA.
AC Q8ZKU9;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
```

```
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Molybdopterin-guanine dinucleotide biosynthesis protein A (Protein
DE FA).
GN Name=moba; OrderedLocusNames=STM3994;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -!- FUNCTION: Links a guanosine 5'-phosphate to molydopterin (MPT)
CC forming molybdopterin guanine dinucleotide (MGD) (By similarity).
CC -!- COFACTOR: Requires magnesium for activity (By similarity).
CC -!- PATHWAY: Molybdenum cofactor biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the mobA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE008886; AAL22833.1; -.
DR HSSP; P32173; 1FR9.
DR StyGene; SG?P?P?; mobA.
DR HAMAP; MF_00316; -.
KW Complete proteome; GTP-binding; Magnesium;
KW Molybdenum cofactor biosynthesis.
SQ SEQUENCE 194 AA; 21189 MW; 09774815EC06F67D CRC64;

Query Match 94.3%; Score 33; DB 1; Length 194;
Best Local Similarity 85.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
Db 14 GGKARRM 20

RESULT 7
O83PF7 PRELIMINARY; PRT; 194 AA.
AC Q83PF7; Q7BZD5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 25-OCT-2004 (TrEMBLrel. 24, Last sequence update)
DE Molybdopterin-guanine dinucleotide biosynthesis protein A.
GN Name=moba; OrderedLocusNames=S3819, S3928;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
```

```
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RL through comparison with genomes of Escherichia coli K12 and O157.";
RN Nucleic Acids Res. 30:4432-4441(2002).
RP [2]
RC SEQUENCE FROM N.A.
RX MEDLINE=22590274; PubMed=12704152;
DOI=10.1128/IAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
Mau B., Berna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RL flexneri serotype 2a strain 2457T.";
DR Infect. Immun. 71:2775-2786(2003).
DR EMBL; AE015400; AAN45363.1; -.
DR EMBL; AE016990; AAP18835.1; -.
DR HSSP; P32173; 1FR9.
KM Complete proteome.
SQ SEQUENCE 194 AA; 21608 MW; 8933CCDEC0DB46BE CRC64;

Query Match
Best Local Similarity 94.3%; Score 33; DB 2; Length 194;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GGKARRL 7
Db 14 GGKARRM 20
```

RESULT 8

```
ID Q13717 PRELIMINARY; PRT; 357 AA.
AC Q13717;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Anion exchange protein 3 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Gehrig H., Appelhans H.;
RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; X70797; CAA50067.1; -.
DR PIR; S31828; S31828.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005452; F:inorganic anion exchanger activity; IEA.
DR GO; GO:0006820; P:anion transport; IEA.
DR InterPro; IPR011531; HCO3_transp.
DR InterPro; IPR003020; HCO3_transp_euk.
DR Pfam; PF00955; HCO3_cotransp; 1.
DR PRINTS; PR00165; ANIONEXCHNGR.
DR PRINTS; PR01231; HCO3TRANSPORT.
DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
FT NON_TER 1
FT NON_TER 357
SQ SEQUENCE 357 AA; 39544 MW; DCB1FA569E19EEF0 CRC64;
```

```
Query Match
Best Local Similarity 94.3%; Score 33; DB 2; Length 357;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GGKARRL 7
Db 138 GGKARRI 144
```

RESULT 9

```
ID Q99416 PRELIMINARY; PRT; 357 AA.
AC Q99416;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Anion exchange protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Gehrig H.H., Koenig J., Appelhans H.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X87211; CAA60670.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005452; F:inorganic anion exchanger activity; IEA.
DR GO; GO:0006820; P:anion transport; IEA.
DR InterPro; IPR001717; Anion_exchange.
DR InterPro; IPR011531; HCO3_transp.
DR InterPro; IPR003020; HCO3_transp_euk.
DR Pfam; PF00955; HCO3_cotransp; 1.
DR PRINTS; PR00165; ANIONEXCHNGR.
DR PRINTS; PR01231; HCO3TRANSPORT.
DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
FT NON_TER 1
FT NON_TER 357
SQ SEQUENCE 357 AA; 39526 MW; 40C03D49F3AF359C CRC64;
```

```
Query Match
Best Local Similarity 94.3%; Score 33; DB 2; Length 357;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GGKARRL 7
Db 138 GGKARRI 144
```

RESULT 10

```
ID Q9ERP4 PRELIMINARY; PRT; 1030 AA.
AC Q9ERP4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Anion exchanger 3 cardiac isoform.
GN Name=Slc4a3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=22481135; PubMed=12592704;
RA Iwasa M., Tatewaki H., Ohno T., Okubo K., Hamasaki N., Kang D.;
RT "Genomic cloning and promoter analysis of a mouse anion exchanger 3
RT (AE3) gene.";
RL DNA Seq. 13:251-255(2002).
DR EMBL; AF294651; AAG25583.1; -.
DR HSSP; P02730; 1BTQ.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005452; F:inorganic anion exchanger activity; IEA.
DR GO; GO:0006820; P:anion transport; IEA.
DR InterPro; IPR001717; Anion_exchange.
DR InterPro; IPR002979; Anion_exngr3.
DR InterPro; IPR011531; HCO3_transp.
DR InterPro; IPR003020; HCO3_transp_euk.
DR Pfam; PF07565; Band_3_cyto; 1.
DR Pfam; PF00955; HCO3_cotransp; 1.
```

```
DR PRINTS; PR00165; ANIONEXCHNGR.
DR PRINTS; PR01189; ANIONEXHNGR3.
DR PRINTS; PR01231; HCO3TRANSPORT.
DR TIGRfams; TIGR00834; ae; 1.
DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
SQ SEQUENCE 1030 AA; 113907 MW; EB3E0BDC0E72576F CRC64;

Query Match          94.3%; Score 33; DB 2; Length 1030;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARRL 7
Db 718 GKGARRI 724

RESULT 11
B3A3_MOUSE
ID B3A3_MOUSE STANDARD; PRT; 1227 AA.
AC P16283;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Anion exchange protein 3 (Neuronal band 3-like protein).
GN Name=Slc4a3; Synonyms=Ae3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90075236; PubMed=2686841; DOI=10.1016/0092-8674(89)90615-6;
RA Kopito R.R., Lee B.S., Simmons D.M., Lindsey A.E., Morgans C.W.,
RA Schneider K.;
RT "Regulation of intracellular pH by a neuronal homolog of the
RT erythrocyte anion exchanger.";
RL Cell 59:927-937(1989).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=94171936; PubMed=8126106;
RA Morgans C.W., Kopito R.R.;
RT "Generation of truncated brain AE3 isoforms by alternate mRNA
RT processing.";
RL J. Cell Sci. 106:1275-1282(1993).
CC -!- FUNCTION: Plasma membrane anion exchange protein.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=Fl-AE3;
CC IsoId=P16283-1; Sequence=Displayed;
CC Name=311-AE3;
CC IsoId=P16283-2; Sequence=VSP_000464, VSP_000465;
CC Name=14-AE3;
CC IsoId=P16283-3; Sequence=VSP_000466, VSP_000467;
CC -!- TISSUE SPECIFICITY: Neuronal.
CC -!- SIMILARITY: Belongs to the anion exchanger family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M28383; AAA37184.1; -.
DR EMBL; S69314; AAB30140.1; -.
DR PIR; A33638; A33638.
DR HSSP; P02730; 1BTQ.
DR MGD; MGI:109350; Slc4a3.
DR InterPro; IPR001717; Anion_exchange.
DR InterPro; IPR002979; Anion_exhngtr3.
```

```
DR InterPro; IPR003020; HCO3_cotransp.
DR Pfam; PF00955; HCO3_cotransp; 1.
DR PRINTS; PR00165; ANIONEXCHNGR.
DR PRINTS; PR01189; ANIONEXHNGR3.
DR PRINTS; PR01231; HCO3TRANSPORT.
DR TIGRfams; TIGR00834; ae; 1.
DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
KW Alternative splicing; Anion exchange; Antiport; Glycoprotein;
KW Ion transport; Lipoprotein; Palmitate; Transmembrane; Transport.
FT DOMAIN 1 707
FT DOMAIN 708 1227 Cytoplasmic.
FT TRANSMEM 708 730 Membrane (anion exchange).
FT TRANSMEM 736 773 Potential.
FT TRANSMEM 793 815 Potential.
FT TRANSMEM 825 846 Potential.
FT TRANSMEM 888 905 Potential.
FT DOMAIN 906 920 Cytoplasmic (Potential).
FT TRANSMEM 921 941 Potential.
FT TRANSMEM 975 997 Potential.
FT TRANSMEM 1023 1044 Potential.
FT TRANSMEM 1078 1123 Potential.
FT TRANSMEM 1150 1186 Potential.
FT DOMAIN 135 145 Poly-Glu.
FT CARBOHYD 868 868 N-linked (GlcNAc...) (Potential).
FT LIPID 1160 1160 S-palmitoyl cysteine (By similarity).
FT VARSPLIC 381 391 AALLDLLEQTL -> RAFWAGNESLL (in isoform
FT 311-AE3).
FT VARSPLIC 392 1227 /FTId=VSP_000464.
FT VARSPLIC 487 503 /FTId=VSP_000465.
FT VARSPLIC 504 1227 /FTId=VSP_000466.
FT VARSPLIC 504 1227 /FTId=VSP_000467.
SQ SEQUENCE 1227 AA; 135164 MW; D5BEC46E03F4251C CRC64;

Query Match          94.3%; Score 33; DB 1; Length 1227;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARRL 7
Db 915 GKGARRI 921

RESULT 12
Q9ERP5 PRELIMINARY; PRT; 1227 AA.
AC Q9ERP5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Anion exchanger 3 brain isoform.
GN Name=Slc4a3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=22481135; PubMed=12592704;
RA Iwasa M., Tatewaki H., Ohno T., Okubo K., Hamasaki N., Kang D.;
RT "Genomic cloning and promoter analysis of a mouse anion exchanger 3
RT (AE3) gene.";
RL DNA Seq. 13:251-255(2002).
DR EMBL; AF294651; AAG25582.1; -.
DR HSSP; P02730; 1BTQ.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005452; F:inorganic anion exchanger activity; IEA.
DR GO; GO:0006820; P:anion transport; IEA.
```



```

DR InterPro; IPR001717; Anion_exchange.
DR InterPro; IPR002979; Anion_exhngtr3.
DR InterPro; IPR011531; HCO3_transpt.
DR InterPro; IPR003020; HCO3_transpt_euk.
DR Pfam; PF07565; Band_3_cyto; 1.
DR Pfam; PF00955; HCO3_cotransp; 1.
DR PRINTS; PRO0165; ANIONEXCHNGR.
DR PRINTS; PRO1189; ANIONEXHNGR3.
DR PRINTS; PRO1231; HCO3TRANSPORT.
DR TIGRFAMs; TIGR00834; ae; 1.
DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
SQ SEQUENCE 1227 AA; 135371 MW; 68BC64B72E5FDC55 CRC64;

Query Match          94.3%; Score 33; DB 2; Length 1227;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARRL 7
Db 915 GKGARRI 921

RESULT 13
B3A3_HUMAN STANDARD; PRT; 1232 AA.
ID B3A3_HUMAN
AC P48751;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Anion exchange protein 3 (Neuronal band 3-like protein) (Cardiac/brain
DE band 3-like protein) (CAE3/BAE3).
GN Name=SLC4A3; Synonyms=AE3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM BAE3).
RC TISSUE=Heart;
RX MEDLINE=95008042; PubMed=7923606;
RA Yannoukakis D., Stuart-Tilley A., Fernandez H., Fey P., Duyk G.,
RA Alper S.L.;
RT "Molecular cloning, expression, and chromosomal localization of two
RT isoforms of the AE3 anion exchanger from human heart.";
RL Circ. Res. 75:603-614(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM CAE3).
RC TISSUE=Heart;
RA Kudo S., Mattei M.-G., Bloor C.M., Fukuda M., Ranney H.M., Xu A.;
RA Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RL
CC -!- FUNCTION: Plasma membrane anion exchange protein of wide
CC distribution. Mediates at least a part of the Cl(-)/HCO3(-)
CC exchange in cardiac myocytes. Both BAE3 and CAE3 forms transport
CC Cl(-).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=BAE3;
CC IsoId=P48751-1; Sequence=Displayed;
CC Name=CAE3;
CC IsoId=P48751-2; Sequence=VSP_000462, VSP_000463;
CC -!- TISSUE SPECIFICITY: Both BAE3 and CAE3 are expressed in failing
CC ventricle.
CC -!- SIMILARITY: Belongs to the anion exchanger family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U05596; AAA50748.1; -.
DR EMBL; L27213; AAB05850.1; -.
DR PIR; I38496; I38496.
DR HSSP; P02730; 1BTQ.
DR Genew; HGNC:11029; SLC4A3.
DR MIM; 106195; -.
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0005624; C:membrane fraction; TAS.
DR GO; GO:0005452; F:inorganic anion exchanger activity; TAS.
DR GO; GO:0006810; P:transport; TAS.
DR InterPro; IPR001717; Anion_exchange.
DR InterPro; IPR002979; Anion_exhngtr3.
DR InterPro; IPR003020; HCO3_cotransp.
DR Pfam; PF00955; HCO3_cotransp; 1.
DR PRINTS; PRO0165; ANIONEXCHNGR.
DR PRINTS; PRO1189; ANIONEXHNGR3.
DR PRINTS; PRO1231; HCO3TRANSPORT.
DR TIGRFAMs; TIGR00834; ae; 1.
DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
DR Alternative splicing; Anion exchange; Antiport; Glycoprotein;
DR Ion transport; Lipoprotein; Palmitate; Transmembrane; Transport.
KW DOMAIN
FT 1 708
FT DOMAIN 1 708
FT TRANSMEM 709 1232
FT TRANSMEM 709 731
FT TRANSMEM 737 774
FT TRANSMEM 794 816
FT TRANSMEM 826 847
FT TRANSMEM 893 910
FT DOMAIN 911 925
FT TRANSMEM 926 946
FT TRANSMEM 980 1002
FT TRANSMEM 1028 1049
FT TRANSMEM 1083 1128
FT TRANSMEM 1155 1191
FT DOMAIN 136 149
FT DOMAIN 152 161
FT DOMAIN 202 207
FT DOMAIN 308 313
FT DOMAIN 442 445
FT DOMAIN 1184 1187
FT LIPID 1165 1165
FT VARSPLIC 1 296
FT VARSPLIC 297 299
FT VARSPLIC 304 304
FT CONFLICT 343 343
FT CONFLICT 466 466
FT CONFLICT 608 608
FT CONFLICT 702 702
FT CONFLICT 812 813
FT CONFLICT 822 822
FT CONFLICT 861 861
FT CONFLICT 863 867
FT CONFLICT 875 875
FT CONFLICT 885 885
FT CONFLICT 887 887
FT CONFLICT 899 899
FT CONFLICT 1096 1096
SQ SEQUENCE 1232 AA; 135706 MW; 1825BE15977C3821 CRC64;

Query Match          94.3%; Score 33; DB 1; Length 1232;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARRL 7
Db 920 GKGARRI 926

```

```
RESULT 14
B3A3_RABIT
ID B3A3_RABIT STANDARD; PRT; 1233 AA.
AC 018917;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Anion exchange protein 3 (Neuronal band 3-like protein) (Anion
  exchanger 3 brain isoform).
GN Name=SLC4A3; Synonyms=AE3;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Abuladze N., Pushkin A., Kurtz I.;
  Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Plasma membrane anion exchange protein.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Neuronal.
CC -1- SIMILARITY: Belongs to the anion exchanger family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use, by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF031650; AAB86859.1; -.
DR HSSP; P02730; 1BTQ.
DR InterPro; IPR001717; Anion_exchange.
DR InterPro; IPR002979; Anion_exhng3.
DR InterPro; IPR003020; HCO3_cotranspt.
DR Pfam; PF00955; HCO3_cotransp; 1.
DR PRINTS; PR00165; ANIONEXCHNGR.
DR PRINTS; PR01189; ANIONEXCHNGR3.
DR PRINTS; PR01231; HCO3TRANSPORT.
DR TIGRfams; TIGR00834; ae; 1.
DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
DR Anion exchange; Antiport; Glycoprotein; Ion transport; Lipoprotein;
  Palmitate; Transmembrane; Transport.
KW DOMAIN
FT 1 709 Cytoplasmic.
FT DOMAIN 1 709 Membrane (anion exchange).
FT TRANSMEM 710 1233 Potential.
FT TRANSMEM 710 732 Potential.
FT TRANSMEM 738 775 Potential.
FT TRANSMEM 795 817 Potential.
FT TRANSMEM 827 848 Potential.
FT TRANSMEM 894 911 Potential.
FT DOMAIN 912 926 Cytoplasmic (Potential).
FT TRANSMEM 927 947 Potential.
FT TRANSMEM 981 1003 Potential.
FT TRANSMEM 1029 1050 Potential.
FT TRANSMEM 1084 1129 Potential.
FT TRANSMEM 1156 1192 Potential.
FT CARBOHYD 874 874 N-linked (GlcNAc...) (Potential).
FT LIPID 1166 1166 S-palmitoyl cysteine (By similarity).
FT DOMAIN 74 83 His-rich.
FT DOMAIN 136 150 Poly-Glu.
FT DOMAIN 203 208 Poly-Ser.
FT DOMAIN 309 312 Poly-Lys.
FT DOMAIN 443 446 Poly-Ser.
FT DOMAIN 1185 1188 Poly-Leu.
SQ SEQUENCE 1233 AA; 135759 MW; AE486423E9818583 CRC64;

Query Match 94.3%; Score 33; DB 1; Length 1233;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GKGARRL 7
Db 921 GKGARRI 927

RESULT 15
O68EG4
ID O68EG4 PRELIMINARY; PRT; 1239 AA.
AC O68EG4;
DT 25-OCT-2004 (TRENDBLREL. 28, Created)
DT 25-OCT-2004 (TRENDBLREL. 28, Last sequence update)
DT 25-OCT-2004 (TRENDBLREL. 28, Last annotation update)
DE SLC4a3 protein.
GN Name=SLC4a3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
  and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC080271; AAH80271.1; -.
DR InterPro; IPR001717; Anion_exchange.
DR InterPro; IPR002979; Anion_exhng3.
DR InterPro; IPR011531; HCO3_transpt.
DR InterPro; IPR003020; HCO3_transpt_euk.
DR Pfam; PF07565; Band_3_cyto; 1.
DR Pfam; PF00955; HCO3_cotransp; 1.
DR PRINTS; PR00165; ANIONEXCHNGR.
DR PRINTS; PR01189; ANIONEXCHNGR3.
DR PRINTS; PR01231; HCO3TRANSPORT.
DR TIGRfams; TIGR00834; ae; 1.
DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
SQ SEQUENCE 1239 AA; 136648 MW; C34DF99759F2C36 CRC64;

Query Match 94.3%; Score 33; DB 2; Length 1239;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARRL 7
Db 915 GKGARRI 921

RESULT 16
O6YIO9
```

ID Q6YIQ9 PRELIMINARY; PRT; 1259 AA.
AC Q6YIQ9;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Anion exchanger SLC4A3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Mount D.B.;
RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY142112; AAN34939.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005452; F:inorganic anion exchanger activity; IEA.
DR GO; GO:0006820; P:anion transport; IEA.
DR InterPro; IPR001717; Anion_exchange.
DR InterPro; IPR002979; Anion_exhngtr3.
DR InterPro; IPR011531; HCO3_transpt.
DR InterPro; IPR003020; HCO3_transpt_euk.
DR Pfam; PF07565; Band_3_cyto; 1.
DR Pfam; PF00955; HCO3_cotransp; 1.
DR PRINTS; PR00165; ANIONEXCHNGR.
DR PRINTS; PR01189; ANIONEXHNGR3.
DR PRINTS; PR01231; HCO3TRANSPORT.
DR TIGRFAMs; TIGR00834; ae; 1.
DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
SQ SEQUENCE 1259 AA; 138654 MW; A145B8B32A318546 CRC64;

Query Match 94.3%; Score 33; DB 2; Length 1259;
Best Local Similarity 85.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
Db 947 GGKARRI 953

RESULT 17
O8YJEB PRELIMINARY; PRT; 67 AA.
ID O8YJEB
AC O8YJEB;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein BMEI0134.
GN OrderedLocusNames=BMEI0134;
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688; DOI=10.1073/pnas.221575398;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mijer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrpides N.C., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AE009456; AAL51316.1; -.
DR PIR; A13268; A13268.
KW Complete proteome.
SQ SEQUENCE 67 AA; 7505 MW; 15399210737D05C2 CRC64;

Query Match 91.4%; Score 32; DB 2; Length 67;
Best Local Similarity 85.7%; Pred. No. 33;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGKARRL 7
Db 12 GGKARRL 18

RESULT 18
O8FYF1 PRELIMINARY; PRT; 67 AA.
ID O8FYF1
AC O8FYF1;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BR1931;
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122; DOI=10.1073/pnas.192319099;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J.A.,
RA Van Aken S.E., Riedmuller S., Tettelin H., Gill S.R., White O.,
RA Salzberg S.L., Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M.,
RA Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
DR EMBL; AE014483; AAN30823.1; -.
DR PIR; A13268; A13268.
DR TIGR; BR1931; -.
KW Complete proteome.
SQ SEQUENCE 67 AA; 7505 MW; 15399210737D05C2 CRC64;

Query Match 91.4%; Score 32; DB 2; Length 67;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
Db 12 GGKARRL 18

RESULT 19
O8GVP5 PRELIMINARY; PRT; 179 AA.
ID O8GVP5
AC O8GVP5;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein P0681F05.121-1.
GN Name=P0681F05.121-1;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, PAC
RT clone:P0681F05.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004674; BAC45150.1; -.
DR Gramene; O8GVP5; -.
KW Hypothetical protein.
SQ SEQUENCE 179 AA; 17722 MW; 46E02AB1C23F334B CRC64;

Query Match 91.4%; Score 32; DB 2; Length 179;
Best Local Similarity 85.7%; Pred. No. 86;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKGARRL 7
Db 121 GKGSRRL 127

RESULT 20
Q62L10 PRELIMINARY; PRT; 234 AA.
AC Q62L10; PRT; 234 AA.
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=BMA0872;
OS Burkholderia mallei ATCC 23344.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=243160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23344;
RA Niernan W.C., Deshazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T., Ulrich R.L., Ronning C.M., Brinkac L.M., Daugherty S.C.,
RA Davidson T.D., Deboy R.T., Dimitrov G., Dodson R.J., Durkin A.S.,
RA Gwinn M.L., Haft D.H., Khouri H., Kolonay J.F., Madupu R.,
RA Mohammad Y., Nelson W.C., Radune D., Romero C.M., Sarria S.,
RA Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y., Zafar N.,
RA Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14247-14251(2004).
DR EMBL; CP000010; AAU49102.1; -.
KW Hypothetical protein.
SQ SEQUENCE 234 AA; 25079 MW; 3EC1D64A6063B812 CRC64;

Query Match 91.4%; Score 32; DB 2; Length 234;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKGARRL 7
Db 30 GGRARRL 36

RESULT 21
Q8L833 PRELIMINARY; PRT; 255 AA.
AC Q8L833;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Putative carbonic anhydrase.
GN Name=At1958180;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Tripp M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY120772; AAM53330.1; -.
DR EMBL; BT000145; AAM15464.1; -.
DR HSSP; P17067; IEKJ.
DR GO; GO:0004089; F:carbonate dehydratase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015976; P:carbon utilization; IEA.
DR InterPro; IPR001765; Prok_plnt_COand.
DR Pfam; PF00484; Pro_CA; 1.
DR PROSITE; PS00704; PROK_CO2_ANHYDRASE_1; 1.
SQ SEQUENCE 255 AA; 29055 MW; C6B6F8AF94612013 CRC64;

Query Match 91.4%; Score 32; DB 2; Length 255;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKGARRL 7
Db 6 GGRARRL 12

RESULT 22
Q9C6R2 PRELIMINARY; PRT; 286 AA.
AC Q9C6R2;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Carbonic anhydrase, putative.
GN Name=TI8124.9;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Niernan W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC079131; AAG50771.1; -.
DR PIR; B96615; B96615.
DR HSSP; P17067; IEKJ.
DR GO; GO:0004089; F:carbonate dehydratase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015976; P:carbon utilization; IEA.
DR InterPro; IPR001765; Prok_plnt_COand.
DR Pfam; PF00484; Pro_CA; 1.
DR PROSITE; PS00704; PROK_CO2_ANHYDRASE_1; 1.
SQ SEQUENCE 286 AA; 32643 MW; 1BAF6DBC194BD61B CRC64;

Query Match 91.4%; Score 32; DB 2; Length 286;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKGARRL 7
Db 6 GGRARRL 12

RESULT 23
Q9C6F5 PRELIMINARY; PRT; 290 AA.
AC Q9C6F5;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)

DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Carbonic anhydrase, putative.
GN Name=T15M6.18;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Roming C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Niernan W.C., Fraser C.M.,
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC079604; AAG50705.1; -.
DR HSSP; P17067; IEKJ.
DR GO; GO:0004089; F:carbonate dehydratase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015976; P:carbon utilization; IEA.
DR InterPro; IPR001765; Prok_plnt_Coanhd.
DR Pfam; PF00484; Pro_CA; 1.
DR PROSITE; PS00704; PROK_CO2_ANHYDRASE 1; 1.
SQ SEQUENCE 290 AA; 33072 MW; 3F63FBA458DEB2A1 CRC64;

Query Match 91.4%; Score 32; DB 2; Length 290;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
Db 6 GGRARRL 12
||:|||||

RESULT 24
Q8TXN6 PRELIMINARY; PRT; 304 AA.
ID Q8TXN6
AC Q8TXN6;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Uncharacterized protein specific for M.kandleri, MK-7 family.
GN OrderedLocusNames=MK0626;
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014; DOI=10.1073/pnas.032671499;
RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malykh A.G., Koonin E.V., Kozyavkin S.A.,
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR EMBL; AE010356; AAM01841.1; -.
KW Complete proteome.
SQ SEQUENCE 304 AA; 33811 MW; 15CC84B674FDF6C CRC64;

Query Match 91.4%; Score 32; DB 2; Length 304;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
Db 296 GGRARRV 302
|||:|||||:

RESULT 25
Q851G5 PRELIMINARY; PRT; 356 AA.
ID Q851G5

AC Q851G5;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein OSUNB0042N1.23.
GN Name=OSUNB0042N1.23;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Ganeberger K., Jones K.M.,
RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC120537; AAO38008.1; -.
DR Gramene; Q851G5; -.
DR InterPro; IPR008976; PLAT_LH2.
KW Hypothetical protein.
SQ SEQUENCE 356 AA; 38198 MW; 84546FB3CD5460E4 CRC64;

Query Match 91.4%; Score 32; DB 2; Length 356;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
Db 87 GGRARRL 93
||:|||||

RESULT 26
Q7WND3 PRELIMINARY; PRT; 416 AA.
ID Q7WND3
AC Q7WND3;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Ferredoxin reductase.
GN OrderedLocusNames=BB1107;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640440; CAE31605.1; -.
DR HSSP; P16640; 1Q1W.
DR GO; GO:0015036; F:disulfide oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.

```
DR InterPro; IPR001327; FAD_pyr_redox.
DR InterPro; IPR000205; NAD_BS.
DR Pfam; PF00070; Pyr_redox; 1.
DR Prodom; PD000139; FAD_pyr_redox; 1.
DR Complete proteome.
SQ SEQUENCE 416 AA; 44039 MW; B95A7FFF3B01FCD6 CRC64;

Query Match
Best Local Similarity 91.4%; Score 32; DB 2; Length 416;
Matches 6; Conservativity 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
Db 118 GGRARRL 124

RESULT 27
O94H89 PRELIMINARY; PRT; 420 AA.
AC Q94H89;
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, last annotation update)
DE Hypothetical protein OSUNBA0090P23.17.
GN Name=OSUNBA0090P23.17;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tseltrin T.,
RA Riggs F., Hsiao J., Zismann V., Blunt S., Pai G., VanAken S.E.,
RA Utterback T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,
RA White O., Fraser C.M.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC084380; AAK52135.1; -.
DR Gramene; Q94H89; -.
KM Hypothetical protein.
SQ SEQUENCE 420 AA; 43986 MW; 6CF074AB7B14B6FC CRC64;

Query Match
Best Local Similarity 91.4%; Score 32; DB 2; Length 420;
Matches 6; Conservativity 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
Db 373 GGRARRL 379

RESULT 28
O45081 PRELIMINARY; PRT; 431 AA.
AC Q45081;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, last annotation update)
DE MOPA.
GN Name=mopA;
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Plasmid pmOP.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pc701;
```

```
RX MEDLINE=96425865; PubMed=8828207;
RA Saint C.P., Romas P.;
RT "4-Methylphthalate catabolism in Burkholderia (Pseudomonas) cepacia
RT Pc701: a gene encoding a phthalate-specific permease forms part of a
RT novel gene cluster.";
RL Microbiology 142:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Pc701;
RA Saint C.P., Boutakis P.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- COFACTOR: FAD (By similarity).
DR EMBL; U29532; AAB41508.1; -.
DR HSSP; Q52437; 1D7Y.
DR GO; GO:0015036; F:disulfide oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001327; FAD_pyr_redox.
DR InterPro; IPR001100; Pyr_redox.
DR Pfam; PF00070; Pyr_redox; 1.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00411; PNDRDTASEI.
DR Prodom; PD000139; FAD_pyr_redox; 1.
KM FAD; Flavoprotein; Oxidoreductase; Plasmid.
SQ SEQUENCE 431 AA; 45627 MW; 331734A2E19AFPA1 CRC64;

Query Match
Best Local Similarity 91.4%; Score 32; DB 2; Length 431;
Matches 6; Conservativity 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
Db 117 GGRARRL 123

RESULT 29
O7WT75 PRELIMINARY; PRT; 487 AA.
AC Q7WT75;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, last annotation update)
DE Phytoene desaturase.
GN Name=crtI;
OS marine bacterium P99-3.
OC Bacteria; Bacteroidetes; Flavobacteria.
OX NCBI_TaxID=216393;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P99-3;
RX MEDLINE=22689040; PubMed=12804761; DOI=10.1016/S0014-5793(03)00513-1;
RA Teramoto M., Takaichi S., Inomata Y., Ikenaga H., Misawa N.;
RT "Structural and functional analysis of a lycopene beta-monocyclase
RT gene isolated from a unique marine bacterium that produces myxol.";
RL FEBS Lett. 545:120-126(2003).
DR EMBL; AB097813; BACT7668.1; -.
DR GO; GO:0050660; F:FAD binding; IEA.
DR GO; GO:0016705; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0016117; P:carotenoid biosynthesis; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002937; Amino_oxidase.
DR InterPro; IPR008151; Phyt_n_dehydro.
DR Pfam; PF01593; Amino_oxidase; 1.
DR Prodom; PD139017; Phyt_n_dehydro; 1.
SQ SEQUENCE 487 AA; 55609 MW; 01C3035BD3BEE39E CRC64;

Query Match
Best Local Similarity 91.4%; Score 32; DB 2; Length 487;
Matches 6; Conservativity 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
Db 38 GGRARRL 44
```

```
RESULT 30
O8QNN6
ID O8QNN6 PRELIMINARY; PRT; 492 AA.
AC O8QNN6;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE EsV-1-25.
GN Name=ORF_25;
OS Ectocarpus siliculosus virus.
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phaeovirus.
OX NCBI_TaxID=37665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EsV-1;
RL Delaroque N., Bothe G., Pohl T., Knippers R., Mueller D.G., Boland W.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF204951; AAK14451.1; -.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR00504; RNA_rec_mot.
DR PROSITE; PS00853; FN3; 1.
DR PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
SQ SEQUENCE 492 AA; 54181 MW; 0536AF583F849902 CRC64;

Query Match          91.4%; Score 32; DB 2; Length 492;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKGARRL 7
Db 478 GGRARRL 484

RESULT 31
O9EWF4
ID O9EWF4 PRELIMINARY; PRT; 548 AA.
AC O9EWF4;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Nitrate reductase beta chain NarH3.
GN ORFNames=2SCK31.08;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939122; CAD30934.1; -.
DR HSSP; P11349; 1R27.
DR GO; GO:0009325; C:nitrate reductase complex; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0008940; F:nitrate reductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0042126; P:nitrate metabolism; IEA.
DR InterPro; IPR001450; 4Fe4S_ferredoxin.
```

```
DR InterPro; IPR000345; CyC heme BS.
DR InterPro; IPR006547; Nit_reduct_beta.
DR PRINTS; PR00353; 4FE4SFRDOXIN.
DR TIGRFAMS; TIGR01660; narH; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 548 AA; 59847 MW; 3A483D2A9B35039B CRC64;

Query Match          91.4%; Score 32; DB 2; Length 548;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKGARRL 7
Db 75 GGRARRL 81

RESULT 32
O65WZ7
ID O65WZ7 PRELIMINARY; PRT; 576 AA.
AC O65WZ7;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Hypothetical protein P0617A08.6.
GN Name=P0617A08.6;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RT "Oryza sativa PAC P0617A08 genomic sequence.";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC135426; AAU44288.1; -.
KW Hypothetical protein.
SQ SEQUENCE 576 AA; 64680 MW; ABAEF97B494A4801 CRC64;

Query Match          91.4%; Score 32; DB 2; Length 576;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKGARRL 7
Db 244 GGRARRL 250

RESULT 33
O92954
ID O92954 PRELIMINARY; PRT; 701 AA.
AC O92954;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Structural polypeptide.
GN Name=gag;
OS Rous sarcoma virus.
OC Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.
OX NCBI_TaxID=11886;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Schmidt-Ruppin B;
RA Bouck J., Skalka A.M., Katz R.A.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to peptidase family A2.
DR EMBL; AF052428; AAC08987.1; -.
```

```
DR PDB; 1D1D; NMR; A=218-479.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0016032; P:viral life cycle; IEA.
DR InterPro; IPR000721; Gag_p24.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR009007; Pept_Aspartic.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR010999; Retrovir_matrix.
DR InterPro; IPR008916; Retrov_capsid_C.
DR InterPro; IPR004028; Retrov_capsid_N.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00607; Gag_p24; 2.
DR Pfam; PF02813; Retro_M; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00098; zf-CCHC; 2.
DR PRINTS; PR00939; C2HCZNFINGER.
DR PRODOM; PD002871; Retro_M; 1.
DR SMART; SM00343; Znf_C2HC; 2.
DR PROSITE; PS00141; A5P_PROTEASE; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS50158; ZF_CCHC; 1.
DR Aspartyl protease; Core protein; Hydrolase; Metal-binding;
KW Polypeptide; Protease; Zinc; Zinc-finger.
FT CHAIN 1 175 p19.
FT CHAIN 178 239 p10.
FT CHAIN 240 479 p27.
FT CHAIN 480 577 p12.
FT CHAIN 578 701 p15.
SQ SEQUENCE 701 AA; 74752 MW; 1004047C574057C7 CRC64;

Query Match          91.4%; Score 32; DB 2; Length 701;
Best Local Similarity 85.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
Db 502 GGRARRL 508

RESULT 34
B3A3_RAT STANDARD; PRT; 1227 AA.
ID B3A3_RAT
AC P23348;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Anion exchange protein 3 (Neuronal band 3-like protein).
GN Name=Slc4a3; Synonyms=Ae3, B3rp3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90094439; PubMed=2294114;
RA Kudrycki K.E., Newman P.R., Shull G.E.;
RT "cDNA cloning and tissue distribution of mRNAs for two proteins that
are related to the band 3 Cl-/HCO3-exchanger.";
RL J. Biol. Chem. 265:462-471(1990).
CC -!- FUNCTION: Plasma membrane anion exchange protein.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Neuronal.
CC -!- SIMILARITY: Belongs to the anion exchanger family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J05167; AAA40798.1; -.
DR PIR; B34911; B34911.
DR HSSP; P02730; 1BTQ.
DR RGD; 3712; Slc4a3.
DR InterPro; IPR001717; Anion_exchange.
DR InterPro; IPR002979; Anion_exhng3.
DR InterPro; IPR003020; HCO3_cotranspt.
DR Pfam; PF00955; HCO3_cotransp; 1.
DR PRINTS; PR00165; ANIONEXCHNGR.
DR PRINTS; PR01189; ANIONEXCHNGR3.
DR PRINTS; PR01231; HCO3TRANSPORT.
DR TIGRfams; TIGR00834; ae; 1.
DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
KW Anion exchange; Antiport; Glycoprotein; Ion transport; Lipoprotein;
KW Palmitate; Transmembrane; Transport.
FT DOMAIN 1 707 Cytoplasmic.
FT DOMAIN 708 1227 Membrane (anion exchange).
FT TRANSMEM 708 730 Potential.
FT TRANSMEM 736 773 Potential.
FT TRANSMEM 793 815 Potential.
FT TRANSMEM 825 846 Potential.
FT TRANSMEM 888 905 Potential.
FT DOMAIN 906 920 Cytoplasmic (Potential).
FT TRANSMEM 921 941 Potential.
FT TRANSMEM 975 997 Potential.
FT TRANSMEM 1023 1044 Potential.
FT TRANSMEM 1078 1123 Potential.
FT TRANSMEM 1150 1186 Potential.
FT CARBOHYD 868 868 N-linked (GlcNAc...) (Potential).
FT LIPID 1160 1160 S-palmitoyl cysteine (By similarity).
FT DOMAIN 74 83 His-rich.
FT DOMAIN 135 149 Poly-Glu.
FT DOMAIN 155 161 Poly-Pro.
FT DOMAIN 307 312 Poly-Lys.
FT DOMAIN 441 444 Poly-Ser.
FT DOMAIN 1179 1182 Poly-Leu.
SQ SEQUENCE 1227 AA; 135406 MW; 3EB1620EE011730E CRC64;

Query Match          91.4%; Score 32; DB 1; Length 1227;
Best Local Similarity 85.7%; Pred. No. 5.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
Db 915 GGRARRV 921

RESULT 35
Q8U2F1 PRELIMINARY; PRT; 48 AA.
ID Q8U2F1
AC Q8U2F1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein PF0885.
GN OrderedlocusNames=PF0885;
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE010203; AAL81009.1; -.
DR GO; GO:0008235; F:metalloexopeptidase activity; IEA.
```


DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
KW Complete proteome.
SQ SEQUENCE 48 AA; 5274 MW; E2EB7762F5E917C4 CRC64;

Query Match 88.6%; Score 31; DB 2; Length 48;
Best Local Similarity 85.7%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGKARRL 7
|||
Db 27 GGGKGRRL 33

RESULT 36

Q6UKB6 PRELIMINARY; PRT; 59 AA.
AC Q6UKB6;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Gp35.
GN Name=ORF35;
OS Burkholderia cepacia phage Bcep43.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=260373;
RN [1]
RP SEQUENCE FROM N.A.
RA Summer E.J., Bomer M.L., Bean E.D., Embry A.E., Mebane L.M.,
RA Tsou L.L.-C., No E.-G., Gonzalez C.F., Young R.F.;
RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY368235; AAR89326.1; -
SQ SEQUENCE 59 AA; 6699 MW; 7C2044DB9F044EDD CRC64;

Query Match 88.6%; Score 31; DB 2; Length 59;
Best Local Similarity 85.7%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGKARRL 7
|||
Db 35 GGGKARRL 41

RESULT 37

Q8HAM5 PRELIMINARY; PRT; 59 AA.
AC Q8HAM5;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein ORF31.
GN Name=ORF31;
OS Burkholderia cepacia phage Bcep781.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=209052;
RN [1]
RP SEQUENCE FROM N.A.
RA Summer E.J., Carlile T.M. Jr., Chastain K.M., Hesteande E.A.,
RA Mebane L.M., Schultz K.E., No E.G., Young R.F., Vidaver A.K.,
RA LiPuma J.J., Gonzalez C.F.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF543311; AAN38035.1; -
KW Hypothetical protein.
SQ SEQUENCE 59 AA; 6699 MW; 7C2044DB9F044EDD CRC64;

Query Match 88.6%; Score 31; DB 2; Length 59;
Best Local Similarity 85.7%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGKARRL 7
|||
Db 35 GGGKARRL 41

RESULT 38
Q6Z463 PRELIMINARY; PRT; 84 AA.
AC Q6Z463;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein P0430F03.45.
GN Name=P0430F03.45;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, PAC
RT clone:P0430F03.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005186; BAC84004.1; -
KW Hypothetical protein.
SQ SEQUENCE 84 AA; 9019 MW; CEC31D71964F3B18 CRC64;

Query Match 88.6%; Score 31; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGKARR 6
|||
Db 23 GGGKARR 28

RESULT 39

H2B1_STRPU STANDARD; PRT; 139 AA.
ID H2B1_STRPU
AC P06145;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Histone H2B.1, sperm.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Sperm;
RX MEDLINE=87066789; PubMed=3786151;
RA Lai Z.-C., Lieber T., Childs G.J.;
RT "The nucleotide sequence of the gene encoding the sperm specific
RT histone subtype H2B-1 from the sea urchin Strongylocentrotus
RT purpuratus.";
RT Nucleic Acids Res. 14:9218-9218(1986).
RL Nucleic Acids Res. 14:9218-9218(1986).
CC -|- SUBUNIT: The nucleosome is an octamer containing two molecules
CC each of H2A, H2B, H3 and H4. The octamer wraps approximately 146
CC bp of DNA.
CC -|- SUBCELLULAR LOCATION: Nuclear.
CC -|- SIMILARITY: Belongs to the histone H2B family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X04681; CAA28385.1; -
DR EMBL; X04681; CAA28384.1; -
DR PIR; S07376; HSURBS.
DR HSSP; P02281; 1KX3.

DR InterPro; IPR007124; Hist_TAF.
DR InterPro; IPR009072; Histone-fold.
DR InterPro; IPR007125; Histone_core_D.
DR InterPro; IPR000558; Histone_H2B.
DR Pfam; PF00125; Histone_1.
DR PRINTS; PR00621; HISTONEH2B.
DR PRODOM; PD000497; Histone_H2B; 1.
DR SMART; SM00427; H2B; 1.
DR PROSITE; PS00357; HISTONE_H2B; 1.
KW Chromosomal protein; DNA-binding; Multigene family; Nuclear protein;
KW Nucleosome core.
FT INIT MET 0
SQ SEQUENCE 139 AA; 15431 MW; C6ED52243709DBDF CRC64;

Query Match
Best Local Similarity 88.6%; Score 31; DB 1; Length 139;
Matches 6; Conservativity 100.0%; Pred. No. 1.1e+02; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARR 6
Db 30 GGKARR 35

RESULT 40
Q8LH66 PRELIMINARY; PRT; 160 AA.
AC Q8LH66;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Hypothetical protein P0668C05.125 (Hypothetical protein OSJNBa0066H10.106).
GN OSJNBa0066H10.106;
OS Name=P0668C05.125; Synonyms=OSJNBa0066H10.106;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa niponbare(GA3) genomic DNA, chromosome 7, PAC clone:P0668C05.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa niponbare(GA3) genomic DNA, chromosome 7, BAC clone:OSJNBa0066H10.";
RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AP004572; BAC10230.1; -.
DR EMBL; AP004572; BAC10230.1; -.
DR EMBL; AP005516; BAD31537.1; -.
DR Gramene; Q8LH66; -.
KW Hypothetical protein.
SQ SEQUENCE 160 AA; 17477 MW; 5E83151103AEC123 CRC64;

Query Match
Best Local Similarity 88.6%; Score 31; DB 2; Length 160;
Matches 6; Conservativity 85.7%; Pred. No. 1.3e+02; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGKARRL 7
Db 62 GGEARRL 68

RESULT 41
Q8TY99 PRELIMINARY; PRT; 165 AA.
AC Q8TY99;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Molybdopterin-guanine dinucleotide biosynthesis protein A.

GN Name=moA; OrderedLocusNames=MK0406;
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014; DOI=10.1073/pnas.032671499;
RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Mal'kh A.G., Koonin E.V., Koz'yavkin S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19 and monophyly of archaeal methanogens."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR EMBL; AE010337; AAM01621.1; -.
KW Complete proteome.
SQ SEQUENCE 165 AA; 18579 MW; 40C051E2DAFB85C4 CRC64;

Query Match
Best Local Similarity 88.6%; Score 31; DB 2; Length 165;
Matches 6; Conservativity 85.7%; Pred. No. 1.3e+02; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGKARRL 7
Db 8 GGKARRL 14

RESULT 42
Q7U8G8 PRELIMINARY; PRT; 173 AA.
AC Q7U8G8;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SYNW0650;
OS Synechococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=84588;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
RA Palenik B., Brahamsha B., Larimer F.W., Land M.L., Hauser L.,
RA Chain P., Lamerdin J.E., Regala W., Allen E.E., McCarron J.,
RA Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synechococcus."
RL Nature 424:1037-1042(2003).
DR EMBL; BX569690; CAE07165.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 173 AA; 19041 MW; 28A2AD916B7D6172 CRC64;

Query Match
Best Local Similarity 88.6%; Score 31; DB 2; Length 173;
Matches 6; Conservativity 100.0%; Pred. No. 1.4e+02; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARR 6
Db 135 GGKARR 140

RESULT 43
Q43589 PRELIMINARY; PRT; 183 AA.
ID Q43589
AC Q43589;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE PAR-1c protein.
GN Name=PAR-1c;
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC lamids; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RX MEDLINE=96145513; PubMed=8555446;
 RA Herbers K., Moenke G., Badur R., Sonnwald U.;
 RT "A simplified procedure for the subtractive cDNA cloning of
 RT photoassimilate-responding genes: isolation of cDNAs encoding a new
 RT class of pathogenesis-related proteins.";
 RL Plant Mol. Biol. 29:1027-1038(1995).
 DR EMBL: X83852; CAA58732.1; -.
 DR PIR: S62700; S62700.
 DR InterPro: IPR009489; PARI.
 DR Pfam: PF06521; PARI; 1.
 SQ SEQUENCE 183 AA; 19857 MW; 62E33FFDB91AF4A1 CRC64;

Query Match 88.6%; Score 31; DB 2; Length 183;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARR 6
 |||||
 Db 146 GGKARR 151

RESULT 44

O9AWL3 PRELIMINARY; PRT; 215 AA.
 AC O9AWL3;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE P0666G04.23 protein.
 GN Name=P0666G04.23;
 OS Oryza sativa (Japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]

RP SEQUENCE FROM N.A.
 RX PubMed=12447438; DOI=10.1038/nature01184;
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
 RA Wu J., Nimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
 RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
 RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
 RA Ikeno M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
 RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
 RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
 RA Yano M., Jiang J., Gojobori T.;
 RT "The genome sequence and structure of rice chromosome 1.";
 RL Nature 420:312-316(2002).
 DR EMBL: AP003047; BAB32921.1; -.
 DR Gramene; O9AWL3; -.
 DR GO: GO:0003677; F:DNA binding; IEA.
 DR InterPro: IPR003657; WRKY.
 DR Pfam: PF03106; WRKY; 1.
 DR PROSITE: PS50811; WRKY; 1.
 SQ SEQUENCE 215 AA; 22268 MW; 87C2A7EB20B63A82 CRC64;

Query Match 88.6%; Score 31; DB 2; Length 215;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARR 6

Db 171 GGKARR 176
 |||||

RESULT 45

O7NKD9 PRELIMINARY; PRT; 246 AA.
 AC O7NKD9;
 DT 01-MAR-2004 (TREMBlrel. 26, Created)
 DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE G11539 protein.
 GN OrderedLocusNames=g11539;
 OS Gloeobacter violaceus.
 OC Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.
 OX NCBI_TaxID=33072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PCC 7421;
 RX MEDLINE=22977040; PubMed=14621292;
 RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
 RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
 RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpo S.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
 RT cyanobacterium that lacks thylakoids.";
 RL DNA Res. 10:137-145(2003).
 DR EMBL: AP006573; BAC89480.1; -.
 KW Complete proteome.
 SQ SEQUENCE 246 AA; 27244 MW; 0A42413CB577E408 CRC64;

Query Match 88.6%; Score 31; DB 2; Length 246;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARR 6
 |||||
 Db 222 GGKARR 227

RESULT 46

O86745 PRELIMINARY; PRT; 258 AA.
 AC O86745;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE ABC transporter ATP-binding protein.
 GN ORFNames=SC6A9.13;
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteridae; Actinobacteriales; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 CC -1- SIMILARITY: Belongs to the ABC transporter family.
 DR EMBL: AL939124; CAA19898.1; -.
 DR PIR: T35444; T35444.
 DR HSSP: Q58206; 1F30.

DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
DR ATP-binding; Complete proteome.
SQ SEQUENCE 258 AA; 27762 MW; 6FD00541723A455D CRC64;

Query Match 88.6%; Score 31; DB 2; Length 258;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARR 6
Db 137 GGKARR 142

RESULT 47

OBLJN4 PRELIMINARY; PRT; 262 AA.
AC OBLJN4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE B1189A09.21 protein.
GN Name=B1189A09.21;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12447438; DOI=10.1038/nature01184;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Ikeda M.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamuya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohka I., Ono N., Saji S., Sakai, K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1."
RL Nature 420:312-316(2002).
DR EMBL; AP003209; BAC00567.1; -.
DR Gramene; OBLJN4; -.
SQ SEQUENCE 262 AA; 28223 MW; 45FC4C64610E5DC5 CRC64;

Query Match 88.6%; Score 31; DB 2; Length 262;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARR 6
Db 243 GGKARR 248

RESULT 48

O8S3R5 PRELIMINARY; PRT; 264 AA.
ID O8S3R5
AC O8S3R5;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Hypothetical protein 49D11.17.
GN Name=49D11.17;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Park Y.-J., Rostoks N., Ramakrishna W., Sanmiguel P., Shiloff B.,
RA Ma J., Jiang Z., Kleinhofs A., Bennettzen J.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF480496; AAL87163.1; -.
DR Gramene; O8S3R5; -.
KW Hypothetical protein.
SQ SEQUENCE 264 AA; 29222 MW; EA5A927544ED5B65 CRC64;

Query Match 88.6%; Score 31; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARR 6
Db 35 GGKARR 40

RESULT 49

O7QER1 PRELIMINARY; PRT; 299 AA.
AC O7QER1;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE AGCP13599 (Fragment).
GN Name=agCG56783; ORFNames=ENSANG00000017421;
OS Anopheles gambiae str. PEST.
OC Anopheles; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB0100846; EAA06563.1; -.
FT NON_TER 1
SQ SEQUENCE 299 AA; 33527 MW; 745815593CDBE5ED CRC64;

Query Match 88.6%; Score 31; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARR 6
Db 173 GGKARR 178

RESULT 50

O6NK28 PRELIMINARY; PRT; 315 AA.
ID O6NK28
AC O6NK28;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, last annotation update)
DE Putative phage protein.
GN OrderedLocustNames=DIP0210;
OS Corynebacterium diptheriae.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1717;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Biotype gravis / NCTC 13129;
RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gkg874;
RA Cerdano-Tarraga A.-M., Efstratiou A., Dover L.G., Holden M.T.G.,
Pallen M.J., Bentley S.D., Beara G.S., Churcher C.M., James K.D.,
De Zeyva A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
Hamlin N., Holroyd S., Jagsels K., Moule S., Quail M.A.,
RA Rabinowitz E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whitehead S., Barrell B.G., Parkhill J.;
RT "The complete genome sequence and analysis of Corynebacterium
RT diphtheriae NCTC13129.";
RL Nucleic Acids Res. 31:6516-6523(2003).
DR EMBL; BX248354; CAE48715.1; -.
KW Complete proteome.
SQ SEQUENCE 315 AA; 33885 MW; F29EB1E4AB8475DB CRC64;

Query Match 88.6%; Score 31; DB 2; Length 315;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARR 6
|||
Db 143 GGKARR 148

Search completed: April 28, 2005, 06:22:09
Job time : 78 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 06:02:40 ; Search time 191 Seconds
(without alignments)
14.174 Million cell updates/sec

Title: US-09-627-600D-135
Perfect score: 35
Sequence: 1 GSKARRL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 65 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	35	100.0	7	4	AAB80729	Aab80729 Human gla
2	35	100.0	7	4	AAB80725	Aab80725 Human gla
3	35	100.0	7	4	AAB80728	Aab80728 Human gla
4	35	100.0	7	4	AAB80711	Aab80711 Human gla
5	35	100.0	8	4	AAB80672	Aab80672 Human gla
6	35	100.0	72	5	ABP02837	Abp02837 Human ORF
7	35	100.0	273	7	ABO73739	Ab073739 Pseudomon
8	33	94.3	812	8	ADH22554	Adh22554 Human tra
9	33	94.3	1153	8	ABM83555	Abm83555 Human dia
10	33	94.3	1198	8	ABM83554	Abm83554 Human dia
11	33	94.3	1207	8	ABM83553	Abm83553 Human dia
12	33	94.3	1225	8	ABM83552	Abm83552 Human dia
13	33	94.3	1232	7	ADE80764	Ade80764 Microsate
14	33	94.3	1234	8	ABM83551	Abm83551 Human dia
15	33	94.3	1259	7	ADJ70768	Adj70768 Human hea
16	32	91.4	7	4	AAB80662	Aab80662 Human gla
17	32	91.4	7	4	AAB80715	Aab80715 Human gla
18	32	91.4	7	4	AAB80699	Aab80699 Human gla
19	32	91.4	7	4	AAB80718	Aab80718 Human gla
20	32	91.4	7	5	AU85849	Aau85849 Prostate-
21	32	91.4	145	3	AAB25193	Aab25193 Eucalyptu
22	32	91.4	290	3	AAg23845	Aag23845 Arabidops
23	32	91.4	290	5	ABB91382	Abb91382 Herbicida
24	32	91.4	304	7	ADM26020	Adm26020 Hyperther
25	32	91.4	317	3	AAg23844	Aag23844 Arabidops

26	32	91.4	1030	2	AAW90261	Aaw90261 A. tigrin
27	31	88.6	7	4	AAB80717	Aab80717 Human gla
28	31	88.6	7	4	AAB80710	Aab80710 Human gla
29	31	88.6	7	4	AAB80722	Aab80722 Human gla
30	31	88.6	50	4	AAU50306	Aau50306 Propionib
31	31	88.6	50	6	ABM46825	Abm46825 Propionib
32	31	88.6	83	2	AAW19270	Aaw19270 Lactobaci
33	31	88.6	91	4	AAg75213	Aag75213 Human col
34	31	88.6	116	4	AAW92531	Aaw92531 Human dig
35	31	88.6	116	4	AAU22568	Aau22568 Novel hum
36	31	88.6	116	7	ADB32408	Adb32408 Human nov
37	31	88.6	141	7	ADF59087	Adf59087 Human pol
38	31	88.6	165	7	ADM25800	Adm25800 Hyperther
39	31	88.6	183	2	AAR99585	Aar99585 Virus-ind
40	31	88.6	186	7	ABO78730	Ab078730 Pseudomon
41	31	88.6	268	7	ABO76862	Ab076862 Pseudomon
42	31	88.6	322	7	ABO82862	Ab082862 Pseudomon
43	31	88.6	364	7	ABO80201	Ab080201 Pseudomon
44	31	88.6	369	7	ABO76787	Ab076787 Pseudomon
45	31	88.6	487	8	ADS29007	Ads29007 Bacterial
46	31	88.6	525	3	AAy96776	Aay96776 Z. mays p
47	31	88.6	608	8	ADN23095	Adn23095 Bacterial
48	31	88.6	610	8	ADN23034	Adn23034 Bacterial
49	31	88.6	808	7	ABO84080	Ab084080 Pseudomon
50	30	85.7	7	4	AAB80706	Aab80706 Human gla
51	30	85.7	7	4	AAB80724	Aab80724 Human gla
52	30	85.7	7	4	AAB80712	Aab80712 Human gla
53	30	85.7	7	4	AAB80720	Aab80720 Human gla
54	30	85.7	7	4	AAB80721	Aab80721 Human gla
55	30	85.7	7	4	AAB80669	Aab80669 Human gla
56	30	85.7	7	4	AAB80723	Aab80723 Human gla
57	30	85.7	7	5	AAU85856	Aau85856 Prostate-
58	30	85.7	8	4	AAU80673	Aau80673 Human gla
59	30	85.7	47	4	AAO11431	Aao11431 Human pol
60	30	85.7	191	4	AAW95052	Aaw95052 Human pro
61	30	85.7	194	4	AAU31452	Aau31452 Novel hum
62	30	85.7	217	7	ABO68164	Ab068164 Pseudomon
63	30	85.7	240	7	ABO67846	Ab067846 Pseudomon
64	30	85.7	251	8	ADR94413	Adr94413 Novel S.
65	30	85.7	265	6	ABU00869	Abu00869 S. pneumo

ALIGNMENTS

RESULT 1	
AAB80729	standard; peptide; 7 AA.
XX	
AC	AAB80729;
XX	
DT	02-MAY-2001 (first entry)
XX	
DE	Human glandular kallikrein 2 substrate peptide #59.
XX	
KW	Cleavage; kallikrein 2; hK2; prodrug.
XX	
OS	Synthetic.
XX	
PN	WO200109165-A2.
XX	
PD	08-FEB-2001.
XX	
PF	28-JUL-2000; 2000WO-US040496.
XX	
PR	29-JUL-1999; 99US-0146316P.
XX	
PA	(UYJO) UNIV JOHNS HOPKINS.
XX	
PI	Denmeade SR, Isaacs JT, Lilja H, Christensen SB;
XX	
DR	WPI; 2001-191450/19.
XX	

PT New peptides containing cleavage sites specifically cleaved by human
PT kallikrein 2, useful for producing prodrugs which treat hk2-producing
PT cell proliferative disorders without exhibiting non-specific toxicity.
XX
PS Example 8; Page 29; 38pp; English.
XX
CC The present invention relates to a peptide comprising an amino acid
CC sequence having a cleavage site specific for an enzyme having a
CC proteolytic activity of human kallikrein 2 (hk2), and which is up to 20
CC amino acids in length. The invention is useful for producing a prodrug
CC which involves linking a drug which contains a primary amine to the
CC peptide, in which the linking of the peptide to the drug inhibits the
CC therapeutic activity of the drug
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 35; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKGARRL 7
 |||||
 1 GKGARRL 7
Db

RESULT 2
AAB80725
ID AAB80725 standard; peptide; 7 AA.
XX
AC AAB80725;
XX
DT 02-MAY-2001 (first entry)
XX
DE Human glandular kallikrein 2 substrate peptide #55.
XX
KW Cleavage; kallikrein 2; hk2; prodrug.
XX
OS Synthetic.
XX
PN WO200109165-A2.
XX
PD 08-FEB-2001.
XX
PF 28-JUL-2000; 2000WO-US040496.
XX
PR 29-JUL-1999; 99US-0146316P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Denmeade SR, Isaacs JT, Lilja H, Christensen SB;
XX
DR WPI; 2001-191450/19.
XX
PT New peptides containing cleavage sites specifically cleaved by human
PT kallikrein 2, useful for producing prodrugs which treat hk2-producing
PT cell proliferative disorders without exhibiting non-specific toxicity.
XX
PS Example 8; Page 29; 38pp; English.
XX
CC The present invention relates to a peptide comprising an amino acid
CC sequence having a cleavage site specific for an enzyme having a
CC proteolytic activity of human kallikrein 2 (hk2), and which is up to 20
CC amino acids in length. The invention is useful for producing a prodrug
CC which involves linking a drug which contains a primary amine to the
CC peptide, in which the linking of the peptide to the drug inhibits the
CC therapeutic activity of the drug
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 35; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKGARRL 7
 |||||
 1 GKGARRL 7
Db

RESULT 3
AAB80728
ID AAB80728 standard; peptide; 7 AA.
XX
AC AAB80728;
XX
DT 02-MAY-2001 (first entry)
XX
DE Human glandular kallikrein 2 substrate peptide #58.
XX
KW Cleavage; kallikrein 2; hk2; prodrug.
XX
OS Synthetic.
XX
PN WO200109165-A2.
XX
PD 08-FEB-2001.
XX
PF 28-JUL-2000; 2000WO-US040496.
XX
PR 29-JUL-1999; 99US-0146316P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Denmeade SR, Isaacs JT, Lilja H, Christensen SB;
XX
DR WPI; 2001-191450/19.
XX
PT New peptides containing cleavage sites specifically cleaved by human
PT kallikrein 2, useful for producing prodrugs which treat hk2-producing
PT cell proliferative disorders without exhibiting non-specific toxicity.
XX
PS Example 8; Page 29; 38pp; English.
XX
CC The present invention relates to a peptide comprising an amino acid
CC sequence having a cleavage site specific for an enzyme having a
CC proteolytic activity of human kallikrein 2 (hk2), and which is up to 20
CC amino acids in length. The invention is useful for producing a prodrug
CC which involves linking a drug which contains a primary amine to the
CC peptide, in which the linking of the peptide to the drug inhibits the
CC therapeutic activity of the drug
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 35; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKGARRL 7
 |||||
 1 GKGARRL 7
Db

RESULT 4
AAB80711
ID AAB80711 standard; peptide; 7 AA.
XX
AC AAB80711;
XX
DT 02-MAY-2001 (first entry)
XX
DE Human glandular kallikrein 2 substrate peptide #41.
XX
KW Cleavage; kallikrein 2; hk2; prodrug.
XX
OS Synthetic.
XX
PN WO200109165-A2.

XX 08-FEB-2001.
PD 28-JUL-2000; 2000WO-US040496.
XX 29-JUL-1999; 99US-0146316P.
XX (UYJO) UNIV JOHNS HOPKINS.
PI Denmeade SR, Isaacs JT, Lilja H, Christensen SB;
XX WPI; 2001-191450/19.
DR
XX
PT New peptides containing cleavage sites specifically cleaved by human
PT kallikrein 2, useful for producing prodrugs which treat hK2-producing
PT cell proliferative disorders without exhibiting non-specific toxicity.
XX
PS Example 8; Page 29; 38pp; English.
XX
SQ Sequence 7 AA;
CC The present invention relates to a peptide comprising an amino acid
CC sequence having a cleavage site specific for an enzyme having a
CC proteolytic activity of human kallikrein 2 (hK2), and which is up to 20
CC amino acids in length. The invention is useful for producing a prodrug
CC which involves linking a drug which contains a primary amine to the
CC peptide, in which the linking of the peptide to the drug inhibits the
CC therapeutic activity of the drug
XX
QY 1 GGKARRL 7
Db 1 GGKARRL 7
RESULTS 5
ID AAB80672 standard; peptide; 8 AA.
XX AAB80672;
AC
XX 02-MAY-2001 (first entry)
DT
XX Human glandular kallikrein 2 substrate peptide #2.
DE
XX Cleavage; kallikrein 2; hK2; prodrug.
KM
XX Synthetic.
OS
XX WO200109165-A2.
PN
XX 08-FEB-2001.
PD
XX 28-JUL-2000; 2000WO-US040496.
PF
XX 29-JUL-1999; 99US-0146316P.
PR
XX (UYJO) UNIV JOHNS HOPKINS.
PA
XX Denmeade SR, Isaacs JT, Lilja H, Christensen SB;
PI WPI; 2001-191450/19.
DR
XX
PT New peptides containing cleavage sites specifically cleaved by human
PT kallikrein 2, useful for producing prodrugs which treat hK2-producing
PT cell proliferative disorders without exhibiting non-specific toxicity.
XX
PS Example 8; Page 28; 38pp; English.
XX
CC The present invention relates to a peptide comprising an amino acid

CC sequence having a cleavage site specific for an enzyme having a
CC proteolytic activity of human kallikrein 2 (hK2), and which is up to 20
CC amino acids in length. The invention is useful for producing a prodrug
CC which involves linking a drug which contains a primary amine to the
CC peptide, in which the linking of the peptide to the drug inhibits the
CC therapeutic activity of the drug
XX
SQ Sequence 8 AA;
CC
QY 1 GGKARRL 7
Db 2 GGKARRL 8
RESULTS 6
ID ABP02837 standard; protein; 72 AA.
XX ABP02837;
AC
XX 24-JUN-2002 (first entry)
DT
XX Human ORFX protein sequence SEQ ID NO:5656.
DE
XX
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KM hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KM degenerative disorder; osteoarthritis; neurodegenerative disorder;
KM cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KM hypertension; hypothyroidism; cholesterol ester storage disease;
KM immune deficiency; immune disorder; infectious disease;
KM autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KM myasthenia gravis.
XX
XX Homo sapiens.
OS
XX WO200192523-A2.
PN
XX 06-DEC-2001.
PD
XX 29-MAY-2001; 2001WO-US010836.
PF
XX 30-MAY-2000; 2000US-0206132P.
PR 29-AUG-2000; 2000US-0228716P.
XX
XX (CURA-) CURAGEN CORP.
PA
XX Shimkets RA, Leach MD;
PI WPI; 2002-106308/14.
DR N-PSDB; ABN18589.
DR
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.
XX
PS Disclosure; SEQ ID NO 5656; 1037pp; English.
XX
CC The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic

CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 72 AA;

Query Match 100.0%; Score 35; DB 5; Length 72;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARRL 7
|||
Db 36 GKGARRL 42

RESULT 7
ABO73739
ID ABO73739 standard; protein; 273 AA.

AC ABO73739;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #5914.

KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

PN US6551795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252391.

PR 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR WPI; 2003-615309/58.

DR N-PSDB; ABD07310.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 22485; 455bp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The

CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX

SQ Sequence 273 AA;

Query Match 100.0%; Score 35; DB 7; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARRL 7
|||
Db 161 GKGARRL 167

RESULT 8
ADH22554
ID ADH22554 standard; protein; 812 AA.

AC ADH22554;

DT 11-MAR-2004 (first entry)

DE Human transporter & ion channel (TRICH) protein SegID52.

KW human; transporters and ion channel; TRICH; cell proliferative;
KW arteriosclerosis; cancer; autoimmune/inflammatory; AIDS; asthma;
KW neurological; epilepsy; stroke; developmental; Cushing's syndrome;
KW hypothyroidism; infection; gene therapy; cytostatic; antiinflammatory;
KW immunosuppressive; antiasthmatic; anticonvulsant; nootropic;
KW neuroprotective; single nucleotide polymorphism; SNP.

XX

OS Homo sapiens.

XX

PN WO2003093444-A2.

PD 13-NOV-2003.

PF 02-MAY-2003; 2003WO-US014026.

PR 03-MAY-2002; 2002US-0377435P.

PR 05-JUN-2002; 2002US-0377444P.

PR 11-JUN-2002; 2002US-0386497P.

PA (INCY-) INCYTE CORP.

PI Baughn MR, Becha SD, Bulloch SA, Chang H, Elliott VS;
PI Emerling BM, Griffin JA, Hafalia AJA, Ison CH, Jackson AA, Jiang X;
PI Jin P, Kable AE, Khare R, Lee SY, Lee S, Mason PM, Marquis JP;
PI Ramkumar J, Richardson TW, Swarnakar A, Tran UK, Chawla NK;
PI Wilson AD;

XX

DR WPI; 2004-022655/02.

DR N-PSDB; ADH22620.

PT New human transporters and ion channels (TRICH), useful for diagnosing,
PT treating and preventing diseases or conditions associated with the
PT aberrant TRICH expression e.g. cancer, AIDS, atherosclerosis, epilepsy,
PT or infections.

XX

PS Claim 1; SEQ ID NO 52; 448bp; English.

XX This invention relates to novel isolated polynucleotides identified as
CC human transporters and ion channels (TRICH), and the encoded polypeptides
CC thereof. Specifically, it describes using these TRICH molecules, as well
CC as agonists, antagonists, antibodies, expression vectors and host cells,
CC in appropriate screening and toxicity assays to assess the effects of
CC exogenous compounds on TRICH expression. The present invention describes
CC TRICH compositions that are useful in the diagnosis, treatment and
CC prevention of various disorders such as cell proliferative (e.g.
CC arteriosclerosis, cancer), autoimmune/inflammatory (e.g. AIDS, asthma),
CC neurological (e.g. epilepsy, stroke) and developmental (e.g. Cushing's

CC syndrome hypothyroidism) and for infections. Accordingly, these TRICH
CC molecules can be used for gene therapy purposes and exhibit various
CC activities such as cytostatic, antiinflammatory, immunosuppressive,
CC antisthmatic, anticonvulsant, nootropic and neuroprotective.
CC Furthermore, a microarray is useful in monitoring or measuring protein-
CC protein interactions, drug-target interactions and gene expression
CC profiles. This polypeptide sequence is a human TRICH protein of the
CC invention.
XX
SQ Sequence 812 AA;

Query Match 94.3%; Score 33; DB 8; Length 812;
Best Local Similarity 85.7%; Pred. No. 6.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
|||||||:
Db 624 GGKARRI 630

RESULT 9
ABM83555
ID ABM83555 standard; protein; 1153 AA.
XX
AC ABM83555;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic pprotein SEQ ID NO:3804.
XX
KW gene therapy; human diagnostic and therapeutic polynucleotide; ditnp.
XX
OS Homo sapiens.
XX
PN WO2004023973-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
PR 12-SEP-2002; 2002US-0410260P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pits SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vilt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX
DR WPI; 2004-329368/30.
DR N-PSDB; ACN42207.
XX
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
PS Claim 27; Page; 190pp; English.
XX
CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (ditnp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The ditnp

CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a ditnp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/1isting.htm
XX
SQ Sequence 1153 AA;

Query Match 94.3%; Score 33; DB 8; Length 1153;
Best Local Similarity 85.7%; Pred. No. 9.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
|||||||:
Db 947 GGKARRI 953

RESULT 10
ABM83554
ID ABM83554 standard; protein; 1198 AA.
XX
AC ABM83554;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic pprotein SEQ ID NO:3803.
XX
KW gene therapy; human diagnostic and therapeutic polynucleotide; ditnp.
XX
OS Homo sapiens.
XX
PN WO2004023973-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
PR 12-SEP-2002; 2002US-0410260P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pits SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vilt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX
DR WPI; 2004-329368/30.
DR N-PSDB; ACN42206.
XX
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
PS Claim 27; Page; 190pp; English.
XX
CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (ditnp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The ditnp
CC molecules may also be used in genetic mapping, in identifying individuals

CC from minute biological samples, in detecting single nucleotide
 CC polymorphisms, as molecular weight markers, and for somatic or germline
 CC gene therapy. The present sequence represents a ditnp protein of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm

XX SQ Sequence 1198 AA;

Query Match 94.3%; Score 33; DB 8; Length 1198;
 Best Local Similarity 85.7%; Pred. No. 9.8e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
 |||||:
 Db 920 GGKARRI 926

RESULT 11

ABM83553 ID ABM83553 standard; protein; 1207 AA.

XX AC ABM83553;

XX DT 18-NOV-2004 (first entry)

XX DE Human diagnostic and therapeutic pprotein SEQ ID NO:3802.

XX KM gene therapy; human diagnostic and therapeutic polynucleotide; ditnp.

XX OS Homo sapiens.

XX PN WO2004023973-A2.

XX PD 25-MAR-2004.

XX PF 12-SEP-2003; 2003WO-US028227.

XX PR 12-SEP-2002; 2002US-0410259P.

XX PR 12-SEP-2002; 2002US-0410260P.

XX PA (INCY-) INCYTE CORP.

XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;

PI Hartschorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;

PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;

PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;

PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;

PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;

PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;

XX PI Patury S, Shi X, Suarez CJ;

XX DR WPI; 2004-329368/30.

XX DR N-PSDB; ACN42205.

XX PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
 PT in diagnosing a condition, disease or disorder associated with human
 PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
 PT in gene mapping.

XX PS Claim 27; Page; 190pp; English.

XX CC The invention relates to novel diagnostic and therapeutic polynucleotides
 CC selected from one of the 2722 sequences defined in the specification. A
 CC polynucleotide of the invention may have a use in gene therapy. The human
 CC diagnostic and therapeutic polynucleotides (ditnp) or polypeptides may be
 CC used to diagnose a particular condition, disease or disorder associated
 CC with human molecules, e.g. cell proliferative disorders,
 CC autoimmune/inflammatory disorder, developmental disorder, endocrine
 CC disorder, neurological disorders, gastrointestinal disorders, or
 CC infections caused by virus, bacteria, fungi or parasite. The ditnp
 CC molecules may also be used in genetic mapping, in identifying individuals
 CC from minute biological samples, in detecting single nucleotide

CC polymorphisms, as molecular weight markers, and for somatic or germline
 CC gene therapy. The present sequence represents a ditnp protein of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm

XX SQ Sequence 1207 AA;

Query Match 94.3%; Score 33; DB 8; Length 1207;
 Best Local Similarity 85.7%; Pred. No. 9.9e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
 |||||:
 Db 920 GGKARRI 926

RESULT 12

ABM83552 ID ABM83552 standard; protein; 1225 AA.

XX AC ABM83552;

XX DT 18-NOV-2004 (first entry)

XX DE Human diagnostic and therapeutic pprotein SEQ ID NO:3801.

XX KM gene therapy; human diagnostic and therapeutic polynucleotide; ditnp.

XX OS Homo sapiens.

XX PN WO2004023973-A2.

XX PD 25-MAR-2004.

XX PF 12-SEP-2003; 2003WO-US028227.

XX PR 12-SEP-2002; 2002US-0410259P.

XX PR 12-SEP-2002; 2002US-0410260P.

XX PA (INCY-) INCYTE CORP.

XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;

PI Hartschorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;

PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;

PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;

PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;

PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;

PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;

XX PI Patury S, Shi X, Suarez CJ;

XX DR WPI; 2004-329368/30.

XX DR N-PSDB; ACN42204.

XX PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
 PT in diagnosing a condition, disease or disorder associated with human
 PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
 PT in gene mapping.

XX PS Claim 27; Page; 190pp; English.

XX CC The invention relates to novel diagnostic and therapeutic polynucleotides
 CC selected from one of the 2722 sequences defined in the specification. A
 CC polynucleotide of the invention may have a use in gene therapy. The human
 CC diagnostic and therapeutic polynucleotides (ditnp) or polypeptides may be
 CC used to diagnose a particular condition, disease or disorder associated
 CC with human molecules, e.g. cell proliferative disorders,
 CC autoimmune/inflammatory disorder, developmental disorder, endocrine
 CC disorder, neurological disorders, gastrointestinal disorders, or
 CC infections caused by virus, bacteria, fungi or parasite. The ditnp
 CC molecules may also be used in genetic mapping, in identifying individuals
 CC from minute biological samples, in detecting single nucleotide
 CC polymorphisms, as molecular weight markers, and for somatic or germline

CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm

XX SQ Sequence 1225 AA;

Query Match 94.3%; Score 33; DB 8; Length 1225;
Best Local Similarity 85.7%; Pred. No. 1e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
|||
Db 947 GGKARRI 953

RESULT 13
ADE80764
ID ADE80764 standard; protein; 1232 AA.
XX
AC ADE80764;
XX
DT 29-JAN-2004 (first entry)
XX
DE Microsatellite related SLC4A3 wt ORF amino acid sequence.
XX
KM frameshift mutation; microsatellite; cytostatic; neuroprotective;
KM vasotrophic; vaccine; gene therapy; neurodegenerative disorder;
KM vascular disease; cancer.
XX
OS Unidentified.
XX
PN WO2003087162-A2.
XX
PD 23-OCT-2003.
XX
PF 17-APR-2003; 2003WO-EP004083.
XX
PR 18-APR-2002; 2002EP-00008771.
PR 18-APR-2002; 2002EP-00008773.
PR 18-APR-2002; 2002EP-00008774.
XX
PA (MTMM-) MTM LAB AG.
XX
PI Von Knebel Doeberitz M, Gebert J, Linnebacher M, Woerner S;
PI Ridder R, Bork P, Yuan YP;
XX
DR WPI; 2003-845308/78.
XX
XX

PT New nucleic acid, useful in preparing a composition for diagnosing or
PT treating disorders associated with frameshift mutations in coding
PT microsatellite regions, e.g., neurodegenerative disorder, vascular
PT disease or cancer.
XX
XX

Claim 3; Fig 2; 62pp; English.

XX The present invention describes a nucleic acid sequence (I) which encodes
CC a polypeptide consisting of TAF1B, MACS, UVRAG, ELAVL3, TCF6L1, ABCF1,
CC AIM2, CHD2, FLJ11053, KIAA1052, ACVR2 or HT001 having a frameshift
CC mutation. Also described: (1) a frameshift polypeptide (II); (2) a method
CC for treating disorders associated with frameshift mutations in coding
CC microsatellites; (3) a pharmaceutical composition comprising the nucleic
CC acid and/or polypeptide; (4) a method for detecting a disorder associated
CC with frameshift mutations in coding microsatellite regions; (5) a
CC diagnostic or research kit for detecting a disorder associated with
CC frameshift mutations in coding microsatellite regions, comprising the
CC nucleic acid and/or frameshift polypeptide; and (6) a method for treating
CC disorders associated with peptides arising from frameshift mutations in
CC coding microsatellite regions in individuals. (I) and (II) have
CC cytostatic, neuroprotective and vasotropic activities, and can be used in
CC vaccines and in gene therapy. The nucleic acid (I) or frameshift
CC polypeptide (II) can be used in detecting disorders associated with
CC frameshift mutations in coding microsatellite regions or in preparing

CC pharmaceutical compositions for treating disorders associated with
CC frameshift mutations in coding microsatellite regions, e.g.,
CC neurodegenerative disorder, vascular disease, cancer or precursory stages
CC of cancer. The present sequence represents a protein used in the
CC exemplification of the present invention.
XX
XX

SQ Sequence 1232 AA;

Query Match 94.3%; Score 33; DB 7; Length 1232;
Best Local Similarity 85.7%; Pred. No. 1e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
|||
Db 920 GGKARRI 926

RESULT 14
ABM83551
ID ABM83551 standard; protein; 1234 AA.
XX
AC ABM83551;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic pprotein SEQ ID NO:3800.
XX
KM gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
OS Homo sapiens.
XX
PN WO2004023973-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
PR 12-SEP-2002; 2002US-0410260P.
XX
PA (INCY-) INCYTE CORP.
XX
XX

PI Schmidt JP, Wright RJ, Bruns CM, Marianovic MM, Shen F;
PI Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerebin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX
XX
DR WPI; 2004-329368/30.
DR N-PSDB; ACN42203.
XX
XX

PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
XX

Claim 27; Page; 190pp; English.

XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline

CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/1listing.htm

XX SQ Sequence 1234 AA;

Query Match 94.3%; Score 33; DB 8; Length 1234;
Best Local Similarity 85.7%; Pred. No. 1e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
|||||:
Db 947 GGKARRI 953

RESULT 15

ADJ70768 standard; protein; 1259 AA.

XX AC ADJ70768;

DT 06-MAY-2004 (first entry)

DE Human heat mitochondrial protein as a therapeutic target SeqID2574.

XX KW mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.

XX OS Homo sapiens.

XX PN WO2003087768-A2.

XX PD 23-OCT-2003.

PF 04-APR-2003; 2003WO-US010870.

PR 12-APR-2002; 2002US-0372843P.

PR 17-JUN-2002; 2002US-0389987P.

PR 20-SEP-2002; 2002US-0412418P.

XX PA (MITO-) MITOKOR.

XX PA (BUCK-) BUCK INST AGE RES.

XX PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;

XX PI Warnock DE;

XX DR WPI; 2003-845369/78.

XX PT Identifying a mitochondrial target for drug screening assays and for

XX PT treating diseases associated with altered mitochondrial function,

XX PT comprises detecting a modified polypeptide in a sample and correlating

XX PT with the disease.

XX PS Claim 1; SEQ ID NO 2574; 180bp; English.

XX CC This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, nootropic, antidiabetic,

CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cyostatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.

XX SQ Sequence 1259 AA;

Query Match 94.3%; Score 33; DB 7; Length 1259;
Best Local Similarity 85.7%; Pred. No. 1e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
|||||:
Db 947 GGKARRI 953

RESULT 16

AAB80662 standard; peptide; 7 AA.

XX AC AAB80662;

DT 02-MAY-2001 (first entry)

DE Human glandular kallikrein 2 cleavage site peptide #23.

XX KW Cleavage; kallikrein 2; hK2; prodrug.

XX OS Synthetic.

XX PN WO200109165-A2.

XX PD 08-FEB-2001.

PF 28-JUL-2000; 2000WO-US040496.

PR 29-JUL-1999; 99US-0146316P.

XX PA (UTJO) UNIV JOHNS HOPKINS.

XX PI Denmeade SR, Isaacs JT, Lilja H, Christensen SB;

XX DR WPI; 2001-191450/19.

XX PT New peptides containing cleavage sites specifically cleaved by human
XX PT kallikrein 2, useful for producing prodrugs which treat hK2-producing
XX PT cell proliferative disorders without exhibiting non-specific toxicity.

XX PS Disclosure; Page 8; 38pp; English.

XX CC The present invention relates to a peptide comprising an amino acid
CC sequence having a cleavage site specific for an enzyme having a
CC proteolytic activity of human kallikrein 2 (hK2), and which is up to 20
CC amino acids in length. The invention is useful for producing a prodrug
CC which involves linking a drug which contains a primary amine to the
CC peptide, in which the linking of the peptide to the drug inhibits the
CC therapeutic activity of the drug

XX SQ Sequence 7 AA;

Query Match 91.4%; Score 32; DB 4; Length 7;
Best Local Similarity 85.7%; Pred. No. 1.8e+06;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
|||||:
Db 1 GGKSRRL 7

RESULT 17

AAB80715 standard; peptide; 7 AA.

XX AC AAB80715;

```
XX 02-MAY-2001 (first entry)
DT
XX
XX Human glandular kallikrein 2 substrate peptide #45.
DE
XX Cleavage; kallikrein 2; hK2; prodrug.
KM
XX
XX Synthetic.
OS
XX WO200109165-A2.
PN
XX 08-FEB-2001.
PD
XX
XX 28-JUL-2000; 2000WO-US040496.
PF
XX 29-JUL-1999; 99US-0146316P.
PR
XX (UYJO ) UNIV JOHNS HOPKINS.
PA
XX
XX Denmeade SR, Isaacs JT, Lilja H, Christensen SB;
PI
XX WPI; 2001-191450/19.
DR
XX
XX New peptides containing cleavage sites specifically cleaved by human
PT kallikrein 2, useful for producing prodrugs which treat hK2-producing
PT cell proliferative disorders without exhibiting non-specific toxicity.
XX
XX Example 8; Page 29; 38pp; English.
PS
XX
XX The present invention relates to a peptide comprising an amino acid
CC sequence having a cleavage site specific for an enzyme having a
CC proteolytic activity of human kallikrein 2 (hK2), and which is up to 20
CC amino acids in length. The invention is useful for producing a prodrug
CC which involves linking a drug which contains a primary amine to the
CC peptide, in which the linking of the peptide to the drug inhibits the
CC therapeutic activity of the drug
CC
XX Sequence 7 AA;
SQ
Query Match 91.4%; Score 32; DB 4; Length 7;
Best Local Similarity 85.7%; Pred. No. 1.8e+06;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGKARRL 7
Db 1 GGKAKRL 7
RESULT 18
AAB80699
ID AAB80699 standard; peptide; 7 AA.
XX
XX AAB80699;
AC
XX
XX 02-MAY-2001 (first entry)
DT
XX
XX Human glandular kallikrein 2 substrate peptide #29.
DE
XX Cleavage; kallikrein 2; hK2; prodrug.
KM
XX
XX Synthetic.
OS
XX
XX WO200109165-A2.
PN
XX
XX 08-FEB-2001.
PD
XX
XX 28-JUL-2000; 2000WO-US040496.
PF
XX 29-JUL-1999; 99US-0146316P.
PR
XX (UYJO ) UNIV JOHNS HOPKINS.
PA
XX
XX Denmeade SR, Isaacs JT, Lilja H, Christensen SB;
PI
```

```
XX WPI; 2001-191450/19.
DR
XX
XX New peptides containing cleavage sites specifically cleaved by human
PT kallikrein 2, useful for producing prodrugs which treat hK2-producing
PT cell proliferative disorders without exhibiting non-specific toxicity.
XX
XX Example 8; Page 29; 38pp; English.
PS
XX
XX The present invention relates to a peptide comprising an amino acid
CC sequence having a cleavage site specific for an enzyme having a
CC proteolytic activity of human kallikrein 2 (hK2), and which is up to 20
CC amino acids in length. The invention is useful for producing a prodrug
CC which involves linking a drug which contains a primary amine to the
CC peptide, in which the linking of the peptide to the drug inhibits the
CC therapeutic activity of the drug
CC
XX Sequence 7 AA;
SQ
Query Match 91.4%; Score 32; DB 4; Length 7;
Best Local Similarity 85.7%; Pred. No. 1.8e+06;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGKARRL 7
Db 1 GGKSRRL 7
RESULT 19
AAB80718
ID AAB80718 standard; peptide; 7 AA.
XX
XX AAB80718;
AC
XX
XX 02-MAY-2001 (first entry)
DT
XX
XX Human glandular kallikrein 2 substrate peptide #48.
DE
XX Cleavage; kallikrein 2; hK2; prodrug.
KM
XX
XX Synthetic.
OS
XX
XX WO200109165-A2.
PN
XX
XX 08-FEB-2001.
PD
XX
XX 28-JUL-2000; 2000WO-US040496.
PF
XX 29-JUL-1999; 99US-0146316P.
PR
XX (UYJO ) UNIV JOHNS HOPKINS.
PA
XX
XX Denmeade SR, Isaacs JT, Lilja H, Christensen SB;
PI
XX WPI; 2001-191450/19.
DR
XX
XX New peptides containing cleavage sites specifically cleaved by human
PT kallikrein 2, useful for producing prodrugs which treat hK2-producing
PT cell proliferative disorders without exhibiting non-specific toxicity.
XX
XX Example 8; Page 29; 38pp; English.
PS
XX
XX The present invention relates to a peptide comprising an amino acid
CC sequence having a cleavage site specific for an enzyme having a
CC proteolytic activity of human kallikrein 2 (hK2), and which is up to 20
CC amino acids in length. The invention is useful for producing a prodrug
CC which involves linking a drug which contains a primary amine to the
CC peptide, in which the linking of the peptide to the drug inhibits the
CC therapeutic activity of the drug
CC
XX Sequence 7 AA;
SQ
Query Match 91.4%; Score 32; DB 4; Length 7;
```

Best Local Similarity 85.7%; Pred. No. 1.8e+06; Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
|||:|
Db 1 GGKARRL 7

RESULT 20
AAU85849
ID AAU85849 standard; peptide; 7 AA.

XX
AC AAU85849;
XX
DT 08-MAY-2002 (first entry)
XX

DE Prostate-specific antigen (PSA)-specific peptide #35.

XX
KW Prostate-specific antigen; PSA; cytostatic; sesquiterpene-gamma-lactone;
KW cell proliferative disorder; malignant cell; non-malignant cell;
KW prostate cancer; thapsigargin.

XX
OS Synthetic.

XX
PN WO200193861-A1.

XX
PD 13-DEC-2001.

XX
PF 28-JUL-2000; 2000WO-US020542.

XX
PR 07-JUN-2000; 2000US-00588822.

XX
PA (UYJO) UNIV JOHNS HOPKINS.

XX
PI Isaacs JT, Lilja H, Denmeade SR, Christensen SB;

XX
DR WPI; 2002-216722/27.

XX
PT New sesquiterpene-gamma-lactone comprising thapsigargin, is useful for
PT treating cell proliferative disorders.

XX
PS Disclosure; Page 18; 72pp; English.

XX
CC The invention relates to a novel therapeutic sesquiterpene-gamma-lactone
CC comprising thapsigargin, and prostate antigen-specific (PSA) peptides.

XX
CC These are used for treating cell proliferative disorders including
CC malignant as well as non-malignant cell populations, such as in prostate
CC cancer. The derivative can be linked to a variety of carrier moieties,

XX
CC such as the peptides of the invention. The linkage converts the
CC derivative into an inactive prodrug. AAU85818-AAU85861 represent the PSA-
CC specific peptides of the invention

XX
SQ Sequence 7 AA;

Query Match 91.4%; Score 32; DB 5; Length 7;
Best Local Similarity 85.7%; Pred. No. 1.8e+06;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
|||:|
Db 1 GGKSRL 7

RESULT 21
AAB25193
ID AAB25193 standard; protein; 145 AA.

XX
AC AAB25193;

XX
DT 27-NOV-2000 (first entry)

XX
DE Eucalyptus grandis cell signalling involved protein SEQ ID NO:512.

KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
KW environmental change; development; cell proliferation; differentiation;
KW elongation; survival; disease resistance; nutrient metabolism.

XX
OS Eucalyptus grandis.

XX
PN WO200042171-A1.

XX
PD 20-JUL-2000.

XX
PF 11-JAN-2000; 2000WO-US000724.

XX
PR 12-JAN-1999; 99US-00228986.

XX
PR 01-NOV-1999; 99US-0162866P.

XX
PA (GENE-) GENESIS RES & DEV CORP LTD.

XX
PI Strabala TJ, Nieuwenhuizen NJ;

XX
DR WPI; 2000-476052/41.

XX
PT Isolated polynucleotide encoding a polypeptide involved in cell signaling
PT used for generating transgenic plants with modified responses to external
PT signals.

XX
PS Claim 3; Page 227; 527pp; English.

XX
CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide

XX
CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or

XX
CC pine (Pinus radiata also known as Monterey pine). The protein sequences

XX
CC are involved in cell signalling. The polynucleotide and protein sequences

XX
CC can be used to modify the response of plant cells to external signals

XX
CC e.g. environmental changes or pathogens during the growth and development

XX
CC of a plant. They can be used to modify cell proliferation,

XX
CC differentiation, elongation and survival, resistance to disease and

XX
CC nutrient metabolism. Examples of modifications which can be produced are

XX
CC altered fruit ripening and senescence of leaves and flowers e.g. to delay

XX
CC senescence and prolong the life of cut flowers or enhance senescence of

XX
CC reproductive organs to engineer sterile plants. Other modifications can

XX
CC be used to delay senescence in selected cell types or organs providing

XX
CC fruit and vegetables which have a longer shelf life between harvest and

XX
CC consumption, or to decrease branching frequency in forest tree species

XX
CC giving long stretches of valuable knot-free clear wood which can be used

XX
CC in solid timber furniture and veneers

SQ Sequence 145 AA;

Query Match 91.4%; Score 32; DB 3; Length 145;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGKARRL 7
|||:|
Db 32 GGKSRL 38

RESULT 22
AAG23845
ID AAG23845 standard; protein; 290 AA.

XX
AC AAG23845;

XX
DT 17-OCT-2000 (first entry)

XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27302.

XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX
OS Arabidopsis thaliana.

PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.

PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148341P.
PR 12-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 13-AUG-1999; 99US-0149368P.
PR 16-AUG-1999; 99US-0149175P.
PR 17-AUG-1999; 99US-0149426P.
PR 18-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.

PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 91.4%; Score 32; DB 3; Length 290;
Best Local Similarity 85.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGRARRL 7
||:||||
Db 6 GGRARRL 12

RESULT 23
ABB91382
ID ABB91382 standard; protein; 290 AA.
XX
AC ABB91382;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 593.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP009892.
XX
PR 28-AUG-2001; 2001WO-EP009892.
XX
PA (FARB) BAYER AG.

XX
PI Tietjen K, Weidler M;
XX
DR WPI, 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.
XX
PS Claim 5; SEQ ID NO 593; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins (ABB90790-ABB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC identifying modulators. The identified modulators are useful as
CC herbicides
XX
SQ Sequence 290 AA;

Query Match 91.4%; Score 32; DB 5; Length 290;
Best Local Similarity 85.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGRARRL 7
||:||||
Db 6 GGRARRL 12

RESULT 24
ADM26020
ID ADM26020 standard; protein; 304 AA.
XX
AC ADM26020;
XX
DT 20-MAY-2004 (first entry)
XX
DE Hyperthermophile Methanopyrus kandleri protein #626.
XX
KW hyperthermophile; protein stability enhancement;
KW protein activity enhancement.
XX
OS Methanopyrus kandleri.
XX
PN WO2003076575-A2.
XX
PD 18-SEP-2003.
XX
PF 04-MAR-2003; 2003WO-US006664.
XX
PR 04-MAR-2002; 2002US-0361742P.
PR 14-MAY-2002; 2002US-0380423P.
PR 16-SEP-2002; 2002US-0410974P.
XX
PA (FIDE-) FIDELITY SYSTEMS INC.
PA (MALY/) MALYKH A.
XX
PI Slesarev AI, Pavlov A, Pavlova N, Kozhavkin S;
XX
DR WPI, 2003-748383/70.
DR N-PSDB; ADM27081.
XX
PT New isolated nucleic acids encoding any of about 1700 Methanopyrus
PT kandleri proteins, and the encoded proteins, useful as a medicaments or
PT as diagnostic agents.
XX
PS Claim 31; SEQ ID NO 626; 1023pp; English.
XX
CC The invention comprises the amino acid sequence of proteins from the

CC hyperthermophile Methanopyrus kandleri, the invention also comprises the
CC complete genome from Methanopyrus kandleri. The Methanopyrus kandleri
CC proteins of the invention are useful for enhancing the stability and/or
CC activity of other proteins. The Methanopyrus kandleri genome is useful in
CC a variety of diagnostic and analytical methods. The present amino acid
CC sequence represents a Methanopyrus kandleri protein of the invention.
XX
SQ Sequence 304 AA;

Query Match 91.4%; Score 32; DB 7; Length 304;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
|||
:|
Db 296 GGKARRV 302

RESULT 25
AAG23844
ID AAG23844 standard; protein; 317 AA.
XX
AC AAG23844;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27301.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.

PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.

PR 29-OCT-1999; 99US-0162142P.

Query Match 91.4%; Score 32; DB 3; Length 317;
Best Local Similarity 85.7%; Pred. No. 4.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKKARRL 7
Db 33 GKKARRL 39

RESULT 26
AAW90261
ID AAW90261 standard; peptide; 1030 AA.
XX
AC AAW90261;
XX
DT 27-APR-1999 (first entry)
XX
DE A. tigrinum AE3 protein.
XX
KW NBC; sodium bicarbonate transporter family; pH regulation; immunoassay;
KW treatment; water retention; blood pressure; acidosis; inflammation;
KW cell proliferation; cancer; sperm activation; inactivation; epilepsy;
KW hydrocephaly; gliucoma; colitis; salamander; AE3; anion exchange.
XX
OS Ambystoma tigrinum.
XX
PN WO9853067-A1.
XX
PD 26-NOV-1998.
XX
PF 20-MAY-1998; 98WO-US010297.
XX
PR 20-MAY-1997; 97US-0047131P.
XX
PA (UYXA) UNIV YALE.
XX
PI Bevensee MO, Schmitt BM, Romero WF, Boron WF, Biemesderfer D;
PI Davis BA, Sussman CR, Choi I, Aalkjaer C, Grichchenko II;
XX
DR WPI; 1999-059743/05.
XX
PT New nucleic acid molecules encoding proteins of the Sodium Bicarbonate
PT Cotransporter (NBC) family - useful for identifying agents that agonise
PT or antagonise NBC activity and treating disorders mediated by NBC.
XX
PS Example 1; Fig 5; 138pp; English.
XX
CC This sequence represents the anion exchange protein AE3 which is used in
CC a method resulting in the isolation of a novel salamander sodium
CC bicarbonate transporter (NBC). NBC proteins and nucleic acid sequences
CC may be used to treat pathological processes including water retention,
CC increased blood pressure, chronic respiratory and metabolic acidosis,
CC inflammation, cell proliferation, cancer, sperm activation/inactivation,
CC hydrocephaly, epilepsy, gliucoma and colitis. Members of the NBC family
CC of proteins can be used (i) as a target to identify agents that block or
CC stimulate NBC mediated pH regulation, (ii) as a target or bait to
CC identify and isolate binding partners that bind NBC, (iii) in methods to
CC identify agents that block or stimulate activity of an NBC protein, and
CC (iv) as a target to assay for NBC-mediated activity. Anti-NBC antibodies
CC are also useful as modulators of NBC activity, useful in the immunoassays
CC for detecting NBC expression/activity and for purifying an NBC protein
XX
SQ Sequence 1030 AA;

Query Match 91.4%; Score 32; DB 2; Length 1030;
Best Local Similarity 85.7%; Pred. No. 1.3e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKKARRL 7
Db 718 GKKARRV 724

RESULT 27
AAB80717
ID AAB80717 standard; peptide; 7 AA.
XX
AC AAB80717;
XX
DT 02-MAY-2001 (first entry)
XX
DE Human glandular kallikrein 2 substrate peptide #47.
XX
KM Cleavage; kallikrein 2; hK2; prodrug.
XX
OS Synthetic.
XX
PN WO200109165-A2.
XX
PD 08-FEB-2001.
XX
PF 28-JUL-2000; 2000WO-US040496.
XX
PR 29-JUL-1999; 99US-0146316P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Denmeade SR, Isaacs JT, Lilja H, Christensen SB;
XX
DR WPI; 2001-191450/19.
XX
PT New peptides containing cleavage sites specifically cleaved by human
PT kallikrein 2, useful for producing prodrugs which treat hK2-producing
PT cell proliferative disorders without exhibiting non-specific toxicity.
XX
PS Example 8; Page 29; 38pp; English.
XX
CC The present invention relates to a peptide comprising an amino acid
CC sequence having a cleavage site specific for an enzyme having a
CC proteolytic activity of human kallikrein 2 (hK2), and which is up to 20
CC amino acids in length. The invention is useful for producing a prodrug
CC which involves linking a drug which contains a primary amine to the
CC peptide, in which the linking of the peptide to the drug inhibits the
CC therapeutic activity of the drug
XX
SQ Sequence 7 AA;

QY 1 GGAARR 6
Db 1 GGAARR 6

Query Match 88.6%; Score 31; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 28
AAB80710
ID AAB80710 standard; peptide; 7 AA.
XX
AC AAB80710;
XX
DT 02-MAY-2001 (first entry)
XX
DE Human glandular kallikrein 2 substrate peptide #40.
XX
KM Cleavage; kallikrein 2; hK2; prodrug.
XX
OS Synthetic.
XX
PN WO200109165-A2.
XX
PD 08-FEB-2001.
XX

PF 28-JUL-2000; 2000WO-US040496.
XX
PR 29-JUL-1999; 99US-0146316P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Denmeade SR, Isaacs JT, Lilja H, Christensen SB;
XX
DR WPI; 2001-191450/19.
XX
PT New peptides containing cleavage sites specifically cleaved by human
PT kallikrein 2, useful for producing prodrugs which treat hK2-producing
PT cell proliferative disorders without exhibiting non-specific toxicity.
XX
PS Example 8; Page 29; 38pp; English.
XX
CC The present invention relates to a peptide comprising an amino acid
CC sequence having a cleavage site specific for an enzyme having a
CC proteolytic activity of human kallikrein 2 (hK2), and which is up to 20
CC amino acids in length. The invention is useful for producing a prodrug
CC which involves linking a drug which contains a primary amine to the
CC peptide, in which the linking of the peptide to the drug inhibits the
CC therapeutic activity of the drug
XX
SQ Sequence 7 AA;

QY 1 GGAARRL 7
Db 1 GGAARRL 7

Query Match 88.6%; Score 31; DB 4; Length 7;
Best Local Similarity 85.7%; Pred. No. 1.8e+06;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 29
AAB80722
ID AAB80722 standard; peptide; 7 AA.
XX
AC AAB80722;
XX
DT 02-MAY-2001 (first entry)
XX
DE Human glandular kallikrein 2 substrate peptide #52.
XX
KM Cleavage; kallikrein 2; hK2; prodrug.
XX
OS Synthetic.
XX
PN WO200109165-A2.
XX
PD 08-FEB-2001.
XX
PF 28-JUL-2000; 2000WO-US040496.
XX
PR 29-JUL-1999; 99US-0146316P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Denmeade SR, Isaacs JT, Lilja H, Christensen SB;
XX
DR WPI; 2001-191450/19.
XX
PT New peptides containing cleavage sites specifically cleaved by human
PT kallikrein 2, useful for producing prodrugs which treat hK2-producing
PT cell proliferative disorders without exhibiting non-specific toxicity.
XX
PS Example 8; Page 29; 38pp; English.
XX
CC The present invention relates to a peptide comprising an amino acid
CC sequence having a cleavage site specific for an enzyme having a
CC proteolytic activity of human kallikrein 2 (hK2), and which is up to 20
CC amino acids in length. The invention is useful for producing a prodrug

CC which involves linking a drug which contains a primary amine to the
CC peptide, in which the linking of the peptide to the drug inhibits the
CC therapeutic activity of the drug
XX
SQ Sequence 7 AA;

Query Match 88.6%; Score 31; DB 4; Length 7;
Best Local Similarity 85.7%; Pred. No. 1.8e+06;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGKARRL 7
Db 1 GGKARQL 7

RESULT 30
AAU50306
ID AAU50306 standard; protein; 50 AA.
XX
AC AAU50306;
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #11202.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59548.
XX

PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
PT

PS Example 1; SEQ ID NO 11501; 1069pp; English.
XX

CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for

CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 50 AA;

Query Match 88.6%; Score 31; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGKARR 6
Db 39 GGKARR 44

RESULT 31
ABM46825
ID ABM46825 standard; protein; 50 AA.
XX
AC ABM46825;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #11501.
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX

XX (CORI-) CORIXA CORP.
XX

PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglass J;
XX

DR WPI; 2003-381789/36.
DR N-PSDB; ACF64477.
XX

PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX

PS Example 1; SEQ ID NO 11501; 1481pp; English.
XX

CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for the
CC nucleic acid hybridisation. The vaccine composition is useful for the

CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 50 AA;

Query Match 88.6%; Score 31; DB 6; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKKARR 6
|||
Db 39 GKKARR 44

RESULT 32
AAW19270

ID AAW19270 standard; protein; 83 AA.

XX
AC AAW19270;

XX 17-OCT-2003 (revised)
DT 15-SEP-1997 (first entry)

XX Lactobacillus amylovorus bacteriocin 1 (Lbnd1).

XX Bacteriocin 1; lactic acid bacterium; vaccine; starter culture; Lbnd1.

XX Lactobacillus amylovorus; LIM KB-180 (LMG P-13139).

XX WO9723619-A1.

XX 03-JUL-1997.

XX 27-NOV-1996; 96WO-EP005235.

XX 22-DEC-1995; 95EP-00870136.

XX (INNO-) INNOGENETICS NV.

XX Contreras B, De Vuyst L, Vandamme E, Sablon E;

XX WPI; 1997-351056/32.

XX N-PSDB; AAT72327.

PT Polynucleic acid encoding Lactobacillus bacteriocins, and immunity and
PT transporter proteins - useful to selectively kill bacteria in therapy and
PT in microbiological and food manufacturing processes.

XX Claim 18; Fig 2; 90pp; English.

CC Bacteriocin 1 (Lbnd1) (AAW19270) is encoded by a gene from operon 5 of a
CC polynucleic acid (AAT72327) isolated from Lactobacillus amylovorus LIM KB
CC -180 (LMG P-13139). Recombinant Lbnd1, and other polypeptides (see also
CC AAW19260-69, AAW19271-72) encoded by the polynucleic acid, can be
CC produced in transformed host cells (e.g. lactic acid bacteria). The
CC transformants may inhibit the growth of other bacteria, or are less
CC sensitive, or resistant, to bacteriocins. Cells that express heterologous
CC immunogens can be used as live oral vaccines, or to express a protein
CC that interferes directly with a toxin. The polypeptide can be used to
CC selectively kill bacteria, useful in therapy and in microbiological and
CC food manufacturing processes. (Updated on 17-OCT-2003 to standardise OS
CC field)

XX Sequence 83 AA;

Query Match 88.6%; Score 31; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKKARR 6
|||
Db 20 GKKARR 25

RESULT 33
AAG75213

ID AAG75213 standard; protein; 91 AA.

XX AAG75213;

XX 03-SEP-2001 (first entry)

XX Human colon cancer antigen protein SEQ ID NO:5977.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma.

XX Homo sapiens.

XX WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US026524.

XX 29-SEP-1999; 99US-0157137P.

XX 03-NOV-1999; 99US-0163280P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

XX N-PSDB; AAH34618.

PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers.

XX Claim 11; Page 7478-7479; 9803pp; English.

CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patients own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922

XX Sequence 91 AA;

Query Match 88.6%; Score 31; DB 4; Length 91;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GKKARRL 7
|||
Db 33 GKKARRL 39

RESULT 34

AAM92531
ID AAM92531 standard; protein; 116 AA.
XX
AC AAM92531;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen SEQ ID NO: 1880.
XX
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum.
XX
OS Homo sapiens.
XX
PN WO200155314-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001324.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.

PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-502630/55.
DR N-PSDB; AAK88304.
XX
PT Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX
PS Claim 11; SEQ ID NO 1880; 986bp; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a digestive system antigen of
CC the invention
XX
SQ Sequence 116 AA;

Query Match 88.6%; Score 31; DB 4; Length 116;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGRARR 6
Db 74 GGRARR 79

RESULT 35
AAU22568
ID AAU22568 standard; protein; 116 AA.
XX
AC AAU22568;
XX
DT 17-DEC-2001 (first entry)
XX
DE Novel human colon associated polypeptide #101.
XX
KW Human; colon cancer; congenital abnormality; infection; colitis;
KW inflammatory bowel disease; IBD; neoplastic disorder; gene therapy;
KW intestinal inflammatory disorder; malabsorption syndrome; gastric;
KW sigmoid disease; antibacterial; antiviral; antiinflammatory; cytostatic.
XX
OS Homo sapiens.
XX
PN WO200155302-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001240.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.

PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465567/50.
DR N-PSDB; AAS39448.
XX
XX
PT Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the colon including colon cancers and also for
PT testing and detection e.g. diagnosis.
XX
XX
PS Claim 11; SEQ ID NO 345; 562pp; English.
XX
CC The present invention relates to the isolation of novel human colon
CC associated polypeptides, and the cDNA (AAS39348-AAS39581) and genomic

CC sequences encoding for them. The sequences of the invention are useful in
CC the diagnosis, treatment, prevention and/or prognosis of disorders of the
CC colon including colon cancer, congenital abnormalities (e.g. atresia and
CC stenosis), bacterial and viral infections, inflammatory bowel disease
CC (IBD), neoplastic cell disorders (e.g. polyps and adenomas, intestinal
CC inflammatory disorders, colitis, colonic inflammation, diarrhoea and
CC dysentery, malabsorption syndromes (e.g. lactose intolerance), intestinal
CC obstruction and sigmoid diseases. The polynucleotide sequences of the
CC invention can also be used in gene therapy. AU22468-AU22701 represent
CC the novel human colon associated polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 116 AA;

Query Match 88.6%; Score 31; DB 4; Length 116;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARR 6
Db 74 GGKARR 79

RESULT 36
ADB32408
ID ADB32408 standard; protein; 116 AA.
XX
AC ADB32408;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human novel colon related polypeptide SEQ ID NO 345.
XX
KW gene therapy; cancer; liver disorder; hepatitis; neural disorder;
KW Alzheimer's disease; human; colon.
XX
OS Homo sapiens.
XX
PN US2003050231-A1.
XX
PD 13-MAR-2003.
XX
PF 17-JAN-2001; 2001US-00764872.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.

PI Rosen CA, Ruben SM, Barash SC;

XX WPI; 2003-625420/59.
XX N-PSDB; ADB32174.

PT New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g. cancer,
PT liver disorders or neural disorders.

PS Claim 12; SEQ ID NO 345; 216pp; English.

XX The invention relates to an isolated nucleic acid molecule encoding a
CC polypeptide. The nucleic acid is useful for preparing a medicament for
CC preventing, treating or ameliorating a medical condition e.g. cancer,
CC liver disorders such as hepatitis or neural disorders such as Alzheimer's
CC disease. The present sequence represents the amino acid sequence of a
CC human novel colon related polypeptide. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030050231.

XX Sequence 116 AA;

Query Match 88.6%; Score 31; DB 7; Length 116;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARR 6
Db 74 GGKARR 79

RESULT 37
ADFS9087
ID ADFS9087 standard; protein; 141 AA.
XX
AC ADFS9087;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human polypeptide sequence SEQ ID NO:1495.
XX
KW biological activity; genetic engineering; hybridisation probe; oligomer;
KW primer; chromosome mapping; gene mapping; recombinant protein production;
KW human.
XX
OS Homo sapiens.
XX
PN WO2003080795-A2.
XX
PD 02-OCT-2003.
XX
PF 09-AUG-2002; 2002WO-US025485.
XX
PR 09-AUG-2001; 2001US-0311261P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Yang Y, Wang Z, Weng G, Ma Y;
XX
DR WPI; 2003-876918/81.
DR N-PSDB; ADFS8087.
XX
PT New polynucleotides, useful as hybridization probes, oligomers or
PT primers, for chromosome or gene mapping, for the recombinant production
PT of proteins, and for generating antisense DNA or RNA.
XX
PS Claim 20; SEQ ID NO 1495; 571pp; English.
XX
CC The present sequence represents a polypeptide (II) with biological
CC activity, which is encoded by an isolated polynucleotide sequence (I);
CC from the present invention. Also described: (1) a vector comprising (I);
CC (2) an expression vector comprising (I); (3) a host cell genetically
CC engineered to comprise (I) which is operatively associated with a
CC regulatory sequence that modulates expression of (I) in the host cell;
CC (4) a polypeptide (II) encoded by (I); (5) a composition comprising the
CC polypeptide of (4) and a carrier; (6) an antibody directed against the
CC polypeptide of (4); (7) detecting (I) or the polypeptide of (4) in a
CC sample; (8) identifying a compound that binds to the polypeptide of (4);
CC (9) producing the polypeptide of (4); and (10) a collection of
CC polynucleotides comprising at least one of the polynucleotide sequences
CC (I). The polynucleotides (I) can be used as hybridisation probes,
CC oligomers or primers, for chromosome or gene mapping, for the recombinant
CC production of proteins, and for generating antisense DNA or RNA.
XX
SQ Sequence 141 AA;

Query Match 88.6%; Score 31; DB 7; Length 141;
Best Local Similarity 85.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
|||:|
Db 94 GGKARQL 100

RESULT 38
ADM25800
ID ADM25800 standard; protein; 165 AA.
XX
AC ADM25800;
XX
DT 20-MAY-2004 (first entry)
XX
DE Hyperthermophile Methanopyrus kandleri protein #406.

XX
KW hyperthermophile; protein stability enhancement;
KW protein activity enhancement.
XX
OS Methanopyrus kandleri.
XX
PN WO2003076575-A2.
XX
PD 18-SEP-2003.
XX
PF 04-MAR-2003; 2003WO-US006664.
XX
PR 04-MAR-2002; 2002US-0361742P.
PR 14-MAY-2002; 2002US-0380423P.
PR 16-SEP-2002; 2002US-0410974P.
XX
PA (FIDE-) FIDELITY SYSTEMS INC.
PA (MALY/) MALYKH A.
XX
PI Slesarev AI, Pavlov A, Pavlova N, Koz'yavkin S;
XX
DR WPI; 2003-748383/70.
DR N-PSDB; ADM27081.
XX
PT New isolated nucleic acids encoding any of about 1700 Methanopyrus
PT kandleri proteins, and the encoded proteins, useful as a medicaments or
PT as diagnostic agents.
XX
PS Claim 31; SEQ ID NO 406; 1023pp; English.
XX
CC The invention comprises the amino acid sequence of proteins from the
CC hyperthermophile Methanopyrus kandleri, the invention also comprises the
CC complete genome from Methanopyrus kandleri. The Methanopyrus kandleri
CC proteins of the invention are useful for enhancing the stability and/or
CC activity of other proteins. The Methanopyrus kandleri genome is useful in
CC a variety of diagnostic and analytical methods. The present amino acid
CC sequence represents a Methanopyrus kandleri protein of the invention.
XX
SQ Sequence 165 AA;

Query Match 88.6%; Score 31; DB 7; Length 165;
Best Local Similarity 85.7%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGKARRL 7
|||:|
Db 8 GGKGRRL 14

RESULT 39
AAR99585
ID AAR99585 standard; protein; 183 AA.
XX
AC AAR99585;
XX
DT 21-NOV-1996 (first entry)
XX
DE Virus-induced protein-20c.
XX
KW Virus-induced protein; VIP; VIP-20a; VIP-20b; VIP-20c; resistance;
KW pathogen.
XX
OS Nicotiana tabacum.
XX
PN DE4442179-A1.
XX
PD 30-MAY-1996.
XX
PF 26-NOV-1994; 94DE-04442179.
XX
PR 26-NOV-1994; 94DE-04442179.
XX
PA (PFLA-) INST PFLANZENGENETIK & KULTURPFLANZENFOR.

XX
PI Sonnewald U, Herbers K;
XX
DR WPI; 1996-260580/27.
DR N-PSDB; AAT32575.
XX
PT DNA encoding virus-induced protein of tobacco - useful for prodn. of
PT transgenic plants having increased pathogen resistance.
XX
PS Disclosure; Fig 4; 21pp; German.
XX
CC VIP-20a, VIP-20b and VIP-20c (AAT32573-T32575) may be used in the
CC transformation of plants to increase resistance against pathogens. The
CC VIPs are isolated from the tobacco plant
XX
SQ Sequence 183 AA;

Query Match 88.6%; Score 31; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARR 6
Db 146 GGKARR 151

RESULT 40
ABO78730
ID ABO78730 standard; protein; 186 AA.
XX
AC ABO78730;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #10905.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
DR N-PSDB; ABD12301.
XX
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 27476; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences

CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 186 AA;

Query Match 88.6%; Score 31; DB 7; Length 186;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARR 6
Db 142 GGKARR 147

RESULT 41
ABO76862
ID ABO76862 standard; protein; 268 AA.
XX
AC ABO76862;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #9037.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
DR N-PSDB; ABD10433.
XX
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 25608; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 268 AA;

Query Match 88.6%; Score 31; DB 7; Length 268;

Best Local Similarity 85.7%; Pred. No. 6.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARRL 7
||:||||
Db 11 GGOARRL 17

RESULT 42
ABO82862
ID ABO82862 standard; protein; 322 AA.

XX ABO82862;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #15037.

KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

PN US6551795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

PT (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR WPI; 2003-615309/58.

DR N-PSDB; ABD16433.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of

PS Disclosure; SEQ ID NO 31608; 455pp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

XX Sequence 322 AA;

Query Match 88.6%; Score 31; DB 7; Length 322;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARR 6
|||||||

Db 317 GKGARR 322

RESULT 43

ABO80201
ID ABO80201 standard; protein; 364 AA.

XX ABO80201;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #12376.

KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

PN US6551795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

PT (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR WPI; 2003-615309/58.

DR N-PSDB; ABD13772.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of

PS Disclosure; SEQ ID NO 28947; 455pp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

XX Sequence 364 AA;

Query Match 88.6%; Score 31; DB 7; Length 364;
Best Local Similarity 100.0%; Pred. No. 8.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARR 6
|||||||

Db 129 GKGARR 134

RESULT 44

ABO76787
ID ABO76787 standard; protein; 369 AA.

XX ABO76787;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #8962.

XX

KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
DR N-PSDB; ABD10358.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 25533; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 369 AA;
OY
1 GGKARRL 7
||| |||
Db 200 GGKTRRL 206
Query Match 88.6%; Score 31; DB 7; Length 369;
Best Local Similarity 85.7%; Pred. No. 8.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
RESULT 45
ADS29007
ID ADS29007 standard; protein; 487 AA.
XX
AC ADS29007;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #18040.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX
OS Bacteria.

XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 18040; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 487 AA;
OY
1 GGKARRL 7
||: ||| |||
Db 178 GGEARRL 184
Query Match 88.6%; Score 31; DB 8; Length 487;
Best Local Similarity 85.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
RESULT 46
AA96776
ID AA96776 standard; protein; 525 AA.
XX
AC AA96776;
XX
DT 09-OCT-2000 (first entry)
XX
DE Z. mays partial weel kinase.
XX
KW Weel; kinase; maize chromosome 4; crop yield; growth; cell cycle;
KW regulator; biomass.

XX OS Zea mays.
XX PN WO200037645-A2.
XX PD 29-JUN-2000.
XX PF 22-DEC-1999; 99WO-US030957.
XX PR 23-DEC-1998; 98US-0113440P.
XX PA (PION-) PIONEER HI-BRED INT INC.
XX PA (UYAR-) UNIV ARIZONA STATE.
XX PI Sun Y, Lowe KS, Dilkes BR, Gordon-Kamm WJ, Larkins BA, Dante RA;
XX DR WPI; 2000-442673/38.
XX DR N-PSDB; AAA51336.
XX PT New maize nucleic acid encoding a cell cycle regulatory protein, known as
XX PT WEE1, useful for modulating the expression of WEE1 polypeptide in plants
XX PT to increase crop yield.
XX PS Disclosure; Page 60-61; 61pp; English.
XX CC The Zea mays weel (zmweel) gene is a single locus on maize chromosome 4
XX CC mapping at the bottom of the long arm at position 187.4 bin 4.11. The
XX CC gene is useful for modulating i.e. for increasing or decreasing the of
XX CC the WEE1 protein in plants such as corn, soybean, sunflower, sorghum,
XX CC canola, wheat, alfalfa, cotton, rice, barley and millet, such that crop
XX CC yield is increased, a positive growth advantage is provided to the plant,
XX CC transformation frequencies which comprise an increase in the number of
XX CC dividing cells to be transformed are improved, endoreduplication which
XX CC involves modulating expression of the WEE1 protein in the endosperm of
XX CC corn, sorghum, wheat, rice, barley and millet are modulated. The weel DNA
XX CC is also used for identifying WEE1 interacting proteins. Modulating cell
XX CC cycle proteins is expected to provide a positive growth advantage, e.g.
XX CC enhance growth and biomass accumulation, and increase crop yield. The
XX CC proteins can be used for identifying compounds that bind to and/or
XX CC increase or decrease the enzymatic activity of catalytically active
XX CC polypeptides
XX SQ Sequence 525 AA;
XX
XX Query Match 88.6%; Score 31; DB 3; Length 525;
XX Best Local Similarity 100.0%; Pred. No. 1.1e+03;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 G GKARR 6
|||
Db 6 G GKARR 11
RESULT 47
ADN23095 ADN23095 standard; protein; 608 AA.
XX
XX AC ADN23095;
XX DT 02-DEC-2004 (first entry)
XX DE Bacterial polypeptide #5748.
XX
XX KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
XX OS Bacteria.
XX

PN US2003233675-A1.
XX
XX PD 18-DEC-2003.
XX PF 20-FEB-2003; 2003US-00369493.
XX PR 21-FEB-2002; 2002US-0360039P.
XX PA (CAOY/) CAO Y.
XX PA (HINK/) HINKLE G J.
XX PA (SLAT/) SLATER S C.
XX PA (CHEN/) CHEN X.
XX PA (GOLD/) GOLDMAN B S.
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX DR WPI; 2004-061375/06.
XX
XX PT New recombinant DNA construct comprising a promoter positioned to provide
XX PT for expression of a polynucleotide encoding a polypeptide from a
XX PT microbial source, useful for producing plants with improved properties.
XX PS Claim 1; SEQ ID NO 5748; 122pp; English.
XX
XX CC The invention relates to a recombinant DNA construct comprising a
XX CC promoter functional in a plant cell, where the promoter is positioned to
XX CC provide for expression of a polynucleotide encoding a polypeptide from a
XX CC microbial source. The invention also relates to a transformed plant
XX CC comprising the recombinant DNA construct and a method of producing a
XX CC transformed plant having an improved property. The plant is a crop plant
XX CC such as maize or soybean. The method of producing a transformed plant
XX CC having an improved property comprises transforming a plant with the
XX CC recombinant DNA construct and growing the transformed plant, where the
XX CC polynucleotide or polypeptide is useful for improving plant properties.
XX CC The recombinant DNA construct is useful for producing plants with
XX CC improved plant properties, e.g. improved cold, heat or drought tolerance,
XX CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX CC increased resistance to plant disease, better growth rate by modification
XX CC of the cell cycle pathway with plant growth regulators, increased rate of
XX CC homologous recombination, modified seed oil or protein yield and/or
XX CC content, improved yield by modification of carbohydrate, nitrogen or
XX CC phosphorus use and/or uptake, by modification of photosynthesis or by
XX CC providing improved plant growth and development under at least one stress
XX CC condition, improved lignin production or improved galactomannan
XX CC production. This sequence represents a bacterial polypeptide used in the
XX CC scope of the invention. Note: The sequence data for this patent did not
XX CC form part of the printed specification but was obtained in electronic
XX CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX SQ Sequence 608 AA;
XX
XX Query Match 88.6%; Score 31; DB 8; Length 608;
XX Best Local Similarity 85.7%; Pred. No. 1.3e+03;
XX Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 G GKARRL 7
||:||||
Db 171 G GEARRL 177
RESULT 48
ADN23034 ADN23034 standard; protein; 610 AA.
XX
XX AC ADN23034;
XX DT 02-DEC-2004 (first entry)
XX DE Bacterial polypeptide #5687.
XX
XX KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;

KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX Bacteria.
XX US2003233675-A1.
XX 18-DEC-2003.
XX 20-FEB-2003; 2003US-00369493.
XX 21-FEB-2002; 2002US-0360039P.
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 5687; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 610 AA;

Query Match 88.6%; Score 31; DB 8; Length 610;
Best Local Similarity 85.7%; Pred. No. 1.3e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGKARRL 7
Db 182 GGEARRL 188

RESULT 49
ABO84080
ID ABO84080 standard; protein; 808 AA.
XX
AC ABO84080;
XX

DT 29-JUL-2004 (first entry)
XX
XX Pseudomonas aeruginosa polypeptide #16255.
DE
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX Pseudomonas aeruginosa.
OS
XX US6551795-B1.
PN
XX 22-APR-2003.
PD
XX 18-FEB-1999; 99US-00252991.
PF
XX 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
PR
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI
XX WPI; 2003-615309/58.
DR
XX N-PSDB; ABD17651.
DR
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 32826; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences ABO67826-
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
SQ Sequence 808 AA;

QY 1 GGGKARR 6
Db 155 GGGKARR 160

RESULT 50
AAB80706
ID AAB80706 standard; peptide; 7 AA.
XX
XX AAB80706;
AC
XX 02-MAY-2001 (first entry)
DT
XX Human glandular kallikrein 2 substrate peptide #36.
DE
XX Cleavage; kallikrein 2; hK2; prodrug.
KW
XX Synthetic.
OS
XX WO200109165-A2.
PN

XX 08-FEB-2001.
PD
XX 28-JUL-2000; 2000WO-US040496.
PF
XX 29-JUL-1999; 99US-0146316P.
PR
XX (UYUO) UNIV JOHNS HOPKINS.
PA
XX Denmeade SR, Isaacs JT, Lilja H, Christensen SB;
PI
XX WPI; 2001-191450/19.
DR
XX
XX New peptides containing cleavage sites specifically cleaved by human
PT kallikrein 2, useful for producing prodrgs which treat hK2-producing
PT cell proliferative disorders without exhibiting non-specific toxicity.
XX
XX Example 8; Page 29; 38pp; English.
PS
XX
XX The present invention relates to a peptide comprising an amino acid
CC sequence having a cleavage site specific for an enzyme having a
CC proteolytic activity of human kallikrein 2 (hK2), and which is up to 20
CC amino acids in length. The invention is useful for producing a prodrg
CC which involves linking a drug which contains a primary amine to the
CC peptide, in which the linking of the peptide to the drug inhibits the
CC therapeutic activity of the drug
XX
SQ Sequence 7 AA;

Query Match 85.7%; Score 30; DB 4; Length 7;
Best Local Similarity 85.7%; Pred. No. 1.8e+06;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GKKARRL 7
||| |||
Db 1 GKKARRL 7

Search completed: April 28, 2005, 06:25:23
Job time : 198 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 06:25:38 ; Search time 134 Seconds
(without alignments)
17.401 Million cell updates/sec

Title: US-09-627-600D-135
Perfect score: 35
Sequence: 1 GGGARRL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1426032 segs, 333106140 residues

Total number of hits satisfying chosen parameters: 1426032

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 65 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	94.3	1259	16 US-10-408-765A-2574	Sequence 2574, Ap
2	32	91.4	145	14 US-10-101-464A-512	Sequence 512, App
3	32	91.4	145	17 US-10-864-252-512	Sequence 512, App
4	32	91.4	179	16 US-10-437-963-137566	Sequence 512, App
5	32	91.4	255	15 US-10-425-114-50835	Sequence 50835, A
6	31	88.6	62	15 US-10-424-599-193894	Sequence 193894,
7	31	88.6	74	16 US-10-437-963-168530	Sequence 168530,
8	31	88.6	91	14 US-10-106-698-5987	Sequence 5987, Ap
9	31	88.6	108	15 US-10-424-599-203122	Sequence 203122,
10	31	88.6	110	16 US-10-437-963-148816	Sequence 148816,
11	31	88.6	116	10 US-09-764-872-345	Sequence 345, App
12	31	88.6	134	15 US-10-424-599-147984	Sequence 147984,
13	31	88.6	163	16 US-10-437-963-117809	Sequence 117809,

14	31	88.6	164	15 US-10-425-114-60348	Sequence 60348, A
15	31	88.6	217	16 US-10-437-963-160663	Sequence 160663,
16	31	88.6	251	16 US-10-767-701-36945	Sequence 36945, A
17	31	88.6	262	16 US-10-437-963-152199	Sequence 152199,
18	31	88.6	313	15 US-10-425-114-50664	Sequence 50664, A
19	31	88.6	443	16 US-10-437-963-104581	Sequence 104581,
20	31	88.6	487	15 US-10-369-493-18040	Sequence 18040, A
21	31	88.6	525	10 US-09-470-526-9	Sequence 9, Appl1
22	31	88.6	608	15 US-10-369-493-5748	Sequence 5748, Ap
23	31	88.6	610	15 US-10-369-493-5687	Sequence 5687, Ap
24	31	88.6	610	16 US-10-437-963-183187	Sequence 183187,
25	31	88.6	724	16 US-10-437-963-162065	Sequence 162065,
26	30	85.7	52	15 US-10-424-599-155164	Sequence 155164,
27	30	85.7	91	15 US-10-424-599-215781	Sequence 215781,
28	30	85.7	157	16 US-10-437-963-148729	Sequence 148729,
29	30	85.7	159	15 US-10-424-599-276484	Sequence 276484,
30	30	85.7	265	17 US-10-472-928-874	Sequence 874, App
31	30	85.7	293	14 US-10-128-714-3296	Sequence 3296, Ap
32	30	85.7	300	15 US-10-425-114-47773	Sequence 47773, A
33	30	85.7	304	15 US-10-425-114-61720	Sequence 61720, A
34	30	85.7	327	14 US-10-128-714-8296	Sequence 8296, Ap
35	30	85.7	634	15 US-10-424-599-190454	Sequence 190454, A
36	30	85.7	641	15 US-10-425-114-50745	Sequence 50745, A
37	30	85.7	651	15 US-10-369-493-21499	Sequence 21499, A
38	30	85.7	1125	16 US-10-437-963-173127	Sequence 173127,
39	29	82.9	55	9 US-09-864-761-47578	Sequence 47578, A
40	29	82.9	68	16 US-10-437-963-124130	Sequence 124130,
41	29	82.9	82	16 US-10-767-701-44489	Sequence 44489, A
42	29	82.9	98	15 US-10-424-599-218288	Sequence 218288,
43	29	82.9	110	16 US-10-437-963-190130	Sequence 190130,
44	29	82.9	120	16 US-10-437-963-187304	Sequence 187304,
45	29	82.9	126	15 US-10-264-049-3223	Sequence 3223, Ap
46	29	82.9	126	16 US-10-767-701-56243	Sequence 56243, A
47	29	82.9	140	16 US-10-767-701-33468	Sequence 33468, A
48	29	82.9	155	15 US-10-424-599-225792	Sequence 225792,
49	29	82.9	169	15 US-10-424-599-220129	Sequence 220129,
50	29	82.9	172	16 US-10-767-701-59569	Sequence 59569, A
51	29	82.9	173	16 US-10-437-963-199628	Sequence 199628,
52	29	82.9	204	15 US-10-425-114-65327	Sequence 65327, A
53	29	82.9	207	16 US-10-437-963-135420	Sequence 135420,
54	29	82.9	218	17 US-10-473-519-23	Sequence 23, Appl
55	29	82.9	239	15 US-10-459-190-13	Sequence 13, Appl
56	29	82.9	264	15 US-10-425-114-53727	Sequence 53727, A
57	29	82.9	269	16 US-10-437-963-170305	Sequence 170305,
58	29	82.9	295	14 US-10-156-761-11899	Sequence 11899, A
59	29	82.9	295	16 US-10-432-934-23	Sequence 23, Appl
60	29	82.9	311	13 US-10-044-442-15	Sequence 15, Appl
61	29	82.9	327	15 US-10-369-493-3029	Sequence 3029, Ap
62	29	82.9	333	15 US-10-282-122A-49692	Sequence 49692, A
63	29	82.9	334	15 US-10-104-047-3786	Sequence 3786, Ap
64	29	82.9	348	15 US-10-425-114-49193	Sequence 49193, A
65	29	82.9	351	15 US-10-424-599-240746	Sequence 240746,

ALIGNMENTS

RESULT 1
US-10-408-765A-2574
; Sequence 2574, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A

```
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2574
; LENGTH: 1259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2574
```

```
Query Match          94.3%; Score 33; DB 16; Length 1259;
Best Local Similarity 85.7%; Pred. No. 7.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGKARRL 7
        |||:|
Db       947 GGKARRI 953
```

RESULT 2

```
US-10-101-464A-512
; Sequence 512, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 512
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-512
```

```
Query Match          91.4%; Score 32; DB 14; Length 145;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGKARRL 7
        |||:|
Db       32 GGKSRRRL 38
```

RESULT 3

```
US-10-864-252-512
; Sequence 512, Application US/10864252
; Publication No. US20050050583A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C3
; CURRENT APPLICATION NUMBER: US/10/864,252
; CURRENT FILING DATE: 2004-06-09
; PRIOR APPLICATION NUMBER: 10/101,464
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
```

```
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 512
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-864-252-512
```

```
Query Match          91.4%; Score 32; DB 17; Length 145;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGKARRL 7
        |||:|
Db       32 GGKSRRRL 38
```

RESULT 4

```
US-10-437-963-137566
; Sequence 137566, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 137566
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_39037C.1.pep
US-10-437-963-137566
```

```
Query Match          91.4%; Score 32; DB 16; Length 179;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGKARRL 7
        |||:|
Db       121 GGKSRRRL 127
```

RESULT 5

```
US-10-425-114-50835
; Sequence 50835, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
```

```
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 50835
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB143-013-H6_FLI.pep
US-10-425-114-50835
```

```
Query Match          91.4%; Score 32; DB 15; Length 255;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GGKARRL 7
         ||:||||
Db       105 GGRARRL 111
```

RESULT 6

```
US-10-424-599-193894
; Sequence 193894, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 193894
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_17111C.1.pep
US-10-424-599-193894
```

```
Query Match          88.6%; Score 31; DB 15; Length 62;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 GGKARRL 7
         ||:||||
Db       1 GKTRRL 7
```

RESULT 7

```
US-10-437-963-168530
; Sequence 168530, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 168530
; LENGTH: 74
```

```
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_67036C.1.pep
US-10-437-963-168530
```

```
Query Match          88.6%; Score 31; DB 16; Length 74;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GGKARRL 7
         ||:||||
Db       50 GGEARRL 56
```

RESULT 8

```
US-10-106-698-5987
; Sequence 5987, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5987
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (52)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (69)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (86)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-5987
```

```
Query Match          88.6%; Score 31; DB 14; Length 91;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 GGKARRL 7
         ||:||||
Db       33 GKGRRRL 39
```

RESULT 9

```
US-10-424-599-203122
; Sequence 203122, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
```

; SEQ ID NO 203122
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_25444C.1.pep
US-10-424-599-203122

Query Match 88.6%; Score 31; DB 15; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARR 6
|||||
Db 29 GKGARR 34

RESULT 10
US-10-437-963-148816
; Sequence 148816, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 148816
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_49209C.1.pep
US-10-437-963-148816

Query Match 88.6%; Score 31; DB 16; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARR 6
|||||
Db 58 GKGARR 63

RESULT 11
US-09-764-872-345
; Sequence 345, Application US/09764872
; Publication No. US20030050231A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA125
; CURRENT APPLICATION NUMBER: US/09/764,872
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 957
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 345
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-872-345

Query Match 88.6%; Score 31; DB 10; Length 116;

Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARR 6
|||||
Db 74 GKGARR 79

RESULT 12
US-10-424-599-147984
; Sequence 147984, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 147984
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(134)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_104651C.1.pep
US-10-424-599-147984

Query Match 88.6%; Score 31; DB 15; Length 134;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARR 6
|||||
Db 11 GKGARR 16

RESULT 13
US-10-437-963-117809
; Sequence 117809, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 117809
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_2117C.1.pep
US-10-437-963-117809

Query Match 88.6%; Score 31; DB 16; Length 163;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGKARRL 7
 ||:||||
 Db 48 GGEARRL 54

RESULT 14
 US-10-425-114-60348
 ; Sequence 60348, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingtong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 60348
 ; LENGTH: 164
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB3607-032-A6_FLI.pep
 US-10-425-114-60348

Query Match 88.6%; Score 31; DB 15; Length 164;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARR 6
 |||||
 Db 3 GGKARR 8

RESULT 15
 US-10-437-963-160663
 ; Sequence 160663, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 160663
 ; LENGTH: 217
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_59922C.1.pep
 US-10-437-963-160663

Query Match 88.6%; Score 31; DB 16; Length 217;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARR 6

Db 156 GGKARR 161
 |||||

RESULT 16
 US-10-767-701-36945
 ; Sequence 36945, Application US/10767701
 ; Publication No. US20040172684A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
 ; FILE REFERENCE: 38-21(53535)B
 ; CURRENT APPLICATION NUMBER: US/10/767,701
 ; CURRENT FILING DATE: 2004-01-29
 ; NUMBER OF SEQ ID NOS: 63128
 ; SEQ ID NO 36945
 ; LENGTH: 251
 ; TYPE: PRT
 ; ORGANISM: Sorghum bicolor
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(251)
 ; OTHER INFORMATION: unsure at all Xaa locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C12737_1.pep
 US-10-767-701-36945

Query Match 88.6%; Score 31; DB 16; Length 251;
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARR 6
 |||||
 Db 20 GGKARR 25

RESULT 17
 US-10-437-963-152199
 ; Sequence 152199, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 152199
 ; LENGTH: 262
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_52274C.1.pep
 US-10-437-963-152199

Query Match 88.6%; Score 31; DB 16; Length 262;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARR 6
 |||||
 Db 243 GGKARR 248

RESULT 18
US-10-425-114-50664
; Sequence 50664, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 50664
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700160276_FLI.pep
US-10-425-114-50664

Query Match 88.6%; Score 31; DB 15; Length 313;
Best Local Similarity 85.7%; Pred. No. 5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKGARRL 7
||:||||
Db 184 GKGARRL 190

RESULT 19
US-10-437-963-104581
; Sequence 104581, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 104581
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_101901C.1.pep
US-10-437-963-104581

Query Match 88.6%; Score 31; DB 16; Length 443;
Best Local Similarity 85.7%; Pred. No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GKGARRL 7
||:||||
Db 238 GKGARRL 244

RESULT 20

US-10-369-493-18040
; Sequence 18040, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18040
; LENGTH: 487
; TYPE: PRT
; ORGANISM: SPHINGOMONAS
US-10-369-493-18040

Query Match 88.6%; Score 31; DB 15; Length 487;
Best Local Similarity 85.7%; Pred. No. 7.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKGARRL 7
||:||||
Db 178 GKGARRL 184

RESULT 21
US-09-470-526-9
; Sequence 9, Application US/09470526
; Publication No. US20030041342A1
; GENERAL INFORMATION:
; APPLICANT: Yuejin Sun
; APPLICANT: Dिल्kes, Brian R.
; APPLICANT: Larkins, Brian A.
; APPLICANT: Lowe, Keith S.
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Dante, Ricardo
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/470,526
; CURRENT FILING DATE: 1999-12-22
; EARLIER APPLICATION NUMBER: 60/109,414
; EARLIER FILING DATE: 1998-12-23
; SOFTWARE: FastSeq for Windows Version 3.0
; NUMBER OF SEQ ID NOS: 9
; SEQ ID NO 9
; LENGTH: 525
; TYPE: PRT
; ORGANISM: zea mays
US-09-470-526-9

Query Match 88.6%; Score 31; DB 10; Length 525;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKGARR 6
||:||||
Db 6 GKGARR 11

RESULT 22
US-10-369-493-5748
; Sequence 5748, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei

```

; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5748
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5748

```

```

Query Match      88.6%; Score 31; DB 15; Length 608;
Best Local Similarity 85.7%; Pred. No. 9.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GGKARRL 7
      ||:||||
Db      171 GGEARRL 177

```

```

RESULT 23
US-10-369-493-5687
; Sequence 5687, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5687
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5687

```

```

Query Match      88.6%; Score 31; DB 15; Length 610;
Best Local Similarity 85.7%; Pred. No. 9.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GGKARRL 7
      ||:||||
Db      182 GGEARRL 188

```

```

RESULT 24
US-10-437-963-183187
; Sequence 183187, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

```

```

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 183187
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_80302C.1.pep
US-10-437-963-183187

```

```

Query Match      88.6%; Score 31; DB 16; Length 610;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GGKARR 6
      ||||||
Db      34 GGKARR 39

```

```

RESULT 25
US-10-437-963-162065
; Sequence 162065, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 162065
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_61191C.1.pep
US-10-437-963-162065

```

```

Query Match      88.6%; Score 31; DB 16; Length 724;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GGKARR 6
      ||||||
Db      76 GGKARR 81

```

```

RESULT 26
US-10-424-599-155164
; Sequence 155164, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28

```

```
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 155164
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_111134C.1.pep
US-10-424-599-155164

Query Match      85.7%; Score 30; DB 15; Length 52;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 G GKARRL 7
        ||| |||
Db      6 GGNARRL 12

RESULT 27
US-10-424-599-215781
; Sequence 215781, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 215781
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_36881C.1.pep
US-10-424-599-215781

Query Match      85.7%; Score 30; DB 15; Length 91;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 G GKARRL 7
        ||| |||
Db      57 G GKLRL 63

RESULT 28
US-10-437-963-148729
; Sequence 148729, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 148729
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Oryza sativa
```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_49130C.1.pep
US-10-437-963-148729

Query Match      85.7%; Score 30; DB 16; Length 157;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 G GKARRL 7
        ||| |||
Db      53 GGSARRL 59

RESULT 29
US-10-424-599-276484
; Sequence 276484, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 276484
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_91685C.1.pep
US-10-424-599-276484

Query Match      85.7%; Score 30; DB 15; Length 159;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 G GKARRL 7
        ||| |||
Db      60 G GKRRL 66

RESULT 30
US-10-472-928-874
; Sequence 874, Application US/10472928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: Seqwin99, version 1.03
; SEQ ID NO 874
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: integrase/recombinase, phage integrase family
; OTHER INFORMATION: Cellular location: cytoplasm
; OTHER INFORMATION: Similar to strain R6 sequence 15902491 (e-154)
US-10-472-928-874

Query Match      85.7%; Score 30; DB 17; Length 265;
Best Local Similarity 85.7%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 1 GGKARRL 7
 ||| |||
 Db 147 GGKRRRL 153

RESULT 31

US-10-128-714-3296
 ; Sequence 3296, Application US/10128714
 ; Publication No. US20030119013A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jiang, Bo
 ; APPLICANT: Hu, Weng
 ; APPLICANT: Tishkoff, Daniel
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Eroshkin, Alexey M
 ; APPLICANT: Lemieux, Sebastien M
 ; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
 ; FILE REFERENCE: 10182-018-999
 ; CURRENT APPLICATION NUMBER: US/10/128,714
 ; PRIOR FILING DATE: 2002-04-23
 ; PRIOR APPLICATION NUMBER: US 60/285,697
 ; PRIOR FILING DATE: 2001-04-23
 ; PRIOR APPLICATION NUMBER: US 60/287,066
 ; PRIOR FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: US 60/295,890
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US 60/303,899
 ; PRIOR FILING DATE: 2001-07-09
 ; PRIOR APPLICATION NUMBER: US 60/316,362
 ; NUMBER OF SEQ ID NOS: 8603
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3296
 ; LENGTH: 293
 ; TYPE: PRT
 ; ORGANISM: Aspergillus fumigatus
 US-10-128-714-3296

Query Match 85.7%; Score 30; DB 14; Length 293;
 Best Local Similarity 85.7%; Pred. No. 7.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGKARRL 7
 ||| |||
 Db 286 GGKRRRL 292

RESULT 32

US-10-425-114-47773
 ; Sequence 47773, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 47773
 ; LENGTH: 300
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB3960-007-G4_FLI.pep
 US-10-425-114-47773

Query Match 85.7%; Score 30; DB 15; Length 300;
 Best Local Similarity 71.4%; Pred. No. 7.6e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
 ||: ||: |
 Db 62 GGRARRI 68

RESULT 33

US-10-425-114-61720
 ; Sequence 61720, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 61720
 ; LENGTH: 304
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: UC-ZMFLB73217C11_FLI.pep
 US-10-425-114-61720

Query Match 85.7%; Score 30; DB 15; Length 304;
 Best Local Similarity 71.4%; Pred. No. 7.7e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
 ||: ||: |
 Db 66 GGRARRI 72

RESULT 34

US-10-128-714-8296
 ; Sequence 8296, Application US/10128714
 ; Publication No. US20030119013A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jiang, Bo
 ; APPLICANT: Hu, Weng
 ; APPLICANT: Tishkoff, Daniel
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Eroshkin, Alexey M
 ; APPLICANT: Lemieux, Sebastien M
 ; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
 ; TITLE OF INVENTION: Methods of Use
 ; FILE REFERENCE: 10182-018-999
 ; CURRENT APPLICATION NUMBER: US/10/128,714
 ; CURRENT FILING DATE: 2002-04-23
 ; PRIOR APPLICATION NUMBER: US 60/285,697
 ; PRIOR FILING DATE: 2001-04-23
 ; PRIOR APPLICATION NUMBER: US 60/287,066
 ; PRIOR FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: US 60/295,890
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US 60/303,899
 ; PRIOR FILING DATE: 2001-07-09
 ; PRIOR APPLICATION NUMBER: US 60/316,362
 ; PRIOR FILING DATE: 2001-08-31
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 8296

; LENGTH: 327
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8296

Query Match 85.7%; Score 30; DB 14; Length 327;
Best Local Similarity 85.7%; Pred. No. 8.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGKARRL 7
|||
Db 320 GGKRRRL 326

RESULT 35
US-10-424-599-190454

; Sequence 190454, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 190454
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(634)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_1429C.1.pep
US-10-424-599-190454

Query Match 85.7%; Score 30; DB 15; Length 634;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGKARRL 7
|||
Db 405 GGKERRL 411

RESULT 36
US-10-425-114-50745

; Sequence 50745, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 50745
; LENGTH: 641
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701145504_FLI.pep

US-10-425-114-50745

Query Match 85.7%; Score 30; DB 15; Length 641;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGKARRL 7
|||
Db 412 GGKERRL 418

RESULT 37

US-10-369-493-21499
; Sequence 21499, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21499
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Methanococcus jannaschii
US-10-369-493-21499

Query Match 85.7%; Score 30; DB 15; Length 651;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGKARRL 7
|||
Db 106 GKANRL 112

RESULT 38

US-10-437-963-173127
; Sequence 173127, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 173127
; LENGTH: 1125
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_71196C.1.pep
US-10-437-963-173127

Query Match 85.7%; Score 30; DB 16; Length 1125;
Best Local Similarity 85.7%; Pred. No. 2.5e+03;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGRARL 7
|||
Db 96 GGSARL 102

RESULT 39

US-09-864-761-47578
; Sequence 47578, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47578
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF191252.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.2
; OTHER INFORMATION: EST_HUMAN HIT: A1161228.1, EVALUUE 1.60e-01
US-09-864-761-47578

Query Match 82.9%; Score 29; DB 9; Length 55;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGRARL 7
|||:|
Db 34 GGSRRV 40

RESULT 40

US-10-437-963-124130
; Sequence 124130, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 124130
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_2689C.1.pep
US-10-437-963-124130

Query Match 82.9%; Score 29; DB 16; Length 68;
Best Local Similarity 71.4%; Pred. No. 3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGRARL 7
||:|
Db 7 GGSRRL 13

RESULT 41

US-10-767-701-44489
; Sequence 44489, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 44489
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB5121-003-A1-PF1-A8.pep
US-10-767-701-44489

Query Match 82.9%; Score 29; DB 16; Length 82;
Best Local Similarity 71.4%; Pred. No. 3.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGRARL 7
||:|
Db 54 GGSRRL 60

RESULT 42

US-10-424-599-218288
; Sequence 218288, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 218288
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(98)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_39141C.1.pep
US-10-424-599-218288

Query Match 82.9%; Score 29; DB 15; Length 98;
Best Local Similarity 71.4%; Pred. No. 4.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
||:|||||
Db 22 GGRARRV 28

RESULT 43

US-10-437-963-190130
; Sequence 190130, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 190130
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(110)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_86573C.1.pep
US-10-437-963-190130

Query Match 82.9%; Score 29; DB 16; Length 110;
Best Local Similarity 71.4%; Pred. No. 4.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7

Db ||:|||||
7 GGRSRRL 13

RESULT 44

US-10-437-963-187304
; Sequence 187304, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 187304
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_8401C.1.pep
US-10-437-963-187304

Query Match 82.9%; Score 29; DB 16; Length 120;
Best Local Similarity 71.4%; Pred. No. 5.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
||:|||||
Db 92 GGRSRRL 98

RESULT 45

US-10-264-049-3223
; Sequence 3223, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3223
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-3223

Query Match 82.9%; Score 29; DB 15; Length 126;
Best Local Similarity 71.4%; Pred. No. 5.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
||:|||||
Db 2 GGRARKL 8

RESULT 46

US-10-767-701-56243

```

; Sequence 56243, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 56243
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30939749.pep
US-10-767-701-56243

```

```

Query Match      82.9%; Score 29; DB 16; Length 126;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2 GKARRL 7
        |||||
Db      117 GKARRL 122

```

```

RESULT 47
US-10-767-701-33468
; Sequence 33468, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 33468
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C32729_1.pep
US-10-767-701-33468

```

```

Query Match      82.9%; Score 29; DB 16; Length 140;
Best Local Similarity 85.7%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 GGAARL 7
        |||||
Db      96 GGAARL 102

```

```

RESULT 48
US-10-424-599-255792
; Sequence 255792, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599

```

```

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 255792
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_73000C.1.pep
US-10-424-599-255792

```

```

Query Match      82.9%; Score 29; DB 15; Length 155;
Best Local Similarity 85.7%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 GGAARL 7
        |||||
Db      133 GGAARL 139

```

```

RESULT 49
US-10-424-599-220129
; Sequence 220129, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 220129
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_40805C.1.pep
US-10-424-599-220129

```

```

Query Match      82.9%; Score 29; DB 15; Length 169;
Best Local Similarity 85.7%; Pred. No. 7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 GGAARL 7
        |||||
Db      50 GGAARL 56

```

```

RESULT 50
US-10-767-701-59569
; Sequence 59569, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 59569
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 7534910.pep
US-10-767-701-59569

```

Query Match 82.9%; Score 29; DB 16; Length 172;
Best Local Similarity 85.7%; Pred. No. 7.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSKARRL 7
| | | | |
Db 64 GSKARRL 70

Search completed: April 28, 2005, 06:41:18
Job time : 135 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 06:06:41 ; Search time 42 Seconds
(without alignments)
12.442 Million cell updates/sec

Title: US-09-627-600D-135
Perfect score: 35
Sequence: 1 GGKARRL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 65 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCtUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	35	100.0	273	4	US-09-252-991A-22485
2	32	91.4	7	4	US-09-627-851B-40
3	32	91.4	145	4	US-10-101-464A-512
4	31	88.6	186	4	US-09-252-991A-27476
5	31	88.6	268	4	US-09-252-991A-25608
6	31	88.6	322	4	US-09-252-991A-31608
7	31	88.6	364	4	US-09-252-991A-28947
8	31	88.6	369	4	US-09-252-991A-25533
9	31	88.6	525	4	US-09-470-526-9
10	31	88.6	808	4	US-09-252-991A-32826
11	30	85.7	7	4	US-09-627-851B-47
12	30	85.7	181	4	US-09-270-767-59934
13	30	85.7	217	4	US-09-252-991A-16910
14	30	85.7	240	4	US-09-252-991A-16592
15	30	85.7	251	4	US-09-107-433-3048
16	30	85.7	686	4	US-09-252-991A-21221
17	29	82.9	204	4	US-09-902-540-16744
18	29	82.9	217	4	US-09-328-352-7258
19	29	82.9	233	4	US-09-252-991A-25021
20	29	82.9	311	3	US-09-087-134-15
21	29	82.9	311	4	US-09-949-016-10383
22	29	82.9	313	4	US-09-252-991A-21867
23	29	82.9	330	4	US-09-252-991A-21479
24	29	82.9	343	4	US-09-252-991A-28970
25	29	82.9	372	4	US-09-252-991A-19721
26	29	82.9	384	4	US-09-552-322-2
27	29	82.9	395	4	US-09-252-991A-17341

28	29	82.9	450	4	US-09-902-540-11536	Sequence 11536, A
29	29	82.9	454	4	US-09-252-991A-22080	Sequence 22080, A
30	29	82.9	458	4	US-09-252-991A-28585	Sequence 28585, A
31	29	82.9	499	4	US-09-561-763-2	Sequence 2, Appli
32	29	82.9	499	4	US-09-431-367B-2	Sequence 2, Appli
33	29	82.9	509	4	US-09-907-794A-315	Sequence 315, App
34	29	82.9	509	4	US-09-905-125A-315	Sequence 315, App
35	29	82.9	509	4	US-09-902-775A-315	Sequence 315, App
36	29	82.9	509	4	US-09-906-700-315	Sequence 315, App
37	29	82.9	509	4	US-09-903-603A-315	Sequence 315, App
38	29	82.9	509	4	US-09-904-920A-315	Sequence 315, App
39	29	82.9	509	4	US-09-909-064-315	Sequence 315, App
40	29	82.9	509	4	US-09-905-381A-315	Sequence 315, App
41	29	82.9	509	4	US-09-906-618-315	Sequence 315, App
42	29	82.9	514	4	US-09-252-991A-28208	Sequence 28208, A
43	29	82.9	544	3	US-09-087-134-14	Sequence 14, Appl
44	29	82.9	579	4	US-09-540-236-2071	Sequence 2071, Ap
45	29	82.9	1413	4	US-09-252-991A-23627	Sequence 23627, A
46	28	80.0	8	2	US-08-504-265B-89	Sequence 89, Appl
47	28	80.0	45	3	US-09-018-635-39	Sequence 39, Appl
48	28	80.0	45	4	US-09-912-962-39	Sequence 39, Appl
49	28	80.0	50	4	US-09-902-540-12424	Sequence 12424, A
50	28	80.0	88	4	US-09-270-767-40821	Sequence 40821, A
51	28	80.0	88	4	US-09-270-767-56037	Sequence 56037, A
52	28	80.0	94	2	US-08-726-306A-167	Sequence 167, App
53	28	80.0	102	4	US-09-252-991A-22830	Sequence 22830, A
54	28	80.0	125	4	US-09-489-039A-9377	Sequence 9377, Ap
55	28	80.0	156	4	US-09-252-991A-32450	Sequence 32450, A
56	28	80.0	168	4	US-09-252-991A-23198	Sequence 23198, A
57	28	80.0	174	4	US-09-252-991A-23538	Sequence 23538, A
58	28	80.0	190	4	US-09-252-991A-18231	Sequence 18231, A
59	28	80.0	197	4	US-09-489-039A-13079	Sequence 13079, A
60	28	80.0	200	4	US-09-949-016-9360	Sequence 9360, Ap
61	28	80.0	215	4	US-09-252-991A-25666	Sequence 25666, A
62	28	80.0	219	4	US-09-252-991A-33077	Sequence 33077, A
63	28	80.0	247	4	US-09-252-991A-29972	Sequence 29972, A
64	28	80.0	248	4	US-09-270-767-36053	Sequence 36053, A
65	28	80.0	248	4	US-09-270-767-51270	Sequence 51270, A

ALIGNMENTS

RESULT 1
US-09-252-991A-22485
; Sequence 22485, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22485
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22485

Query Match 100.0%; Score 35; DB 4; Length 273;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
Db 161 GGKARRL 167

```
RESULT 2
US-09-627-851B-40
; Sequence 40, Application US/09627851B
; Patent No. 6545131
; GENERAL INFORMATION:
; APPLICANT: ISSACS, JOHN T.
; APPLICANT: DENMEADE, SAMUEL R.
; APPLICANT: CHRISTENSEN, S. BROGGER
; APPLICANT: LILJA, HANS
; TITLE OF INVENTION: TISSUE SPECIFIC PRODRUG
; FILE REFERENCE: 57111 (71699)
; CURRENT APPLICATION NUMBER: US/09/627,851B
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/588,822
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: 09/081,707
; PRIOR FILING DATE: 1998-05-19
; PRIOR APPLICATION NUMBER: 60/047,070
; PRIOR FILING DATE: 1997-05-19
; PRIOR APPLICATION NUMBER: 60/080,046
; PRIOR FILING DATE: 1998-03-30
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-627-851B-40
```

```
Query Match      91.4%; Score 32; DB 4; Length 7;
Best Local Similarity 85.7%; Pred. No. 4.1e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGKARRL 7
Db      1 GGKSRRLL 7
```

```
RESULT 3
US-10-101-464A-512
; Sequence 512, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 512
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-512
```

```
Query Match      91.4%; Score 32; DB 4; Length 145;
Best Local Similarity 85.7%; Pred. No. 47;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGKARRL 7
```

```
Db      32 GGKSRRLL 38
```

```
RESULT 4
US-09-252-991A-27476
; Sequence 27476, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27476
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27476
```

```
Query Match      88.6%; Score 31; DB 4; Length 186;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGKARR 6
Db      142 GGKARR 147
```

```
RESULT 5
US-09-252-991A-25608
; Sequence 25608, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25608
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25608
```

```
Query Match      88.6%; Score 31; DB 4; Length 268;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGKARRL 7
Db      11 GGOARRL 17
```

```
RESULT 6
US-09-252-991A-31608
; Sequence 31608, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
```

```
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31608
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31608

Query Match      88.6%; Score 31; DB 4; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GKGARR 6
Db      317 GKGARR 322

RESULT 7
US-09-252-991A-28947
; Sequence 28947, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28947
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28947

Query Match      88.6%; Score 31; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GKGARR 6
Db      129 GKGARR 134

RESULT 8
US-09-252-991A-25533
; Sequence 25533, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25533
; LENGTH: 369
```

```
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25533

Query Match      88.6%; Score 31; DB 4; Length 369;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GKGARRL 7
Db      200 GKGTRRL 206

RESULT 9
US-09-470-526-9
; Sequence 9, Application US/09470526
; Patent No. 6777590
; GENERAL INFORMATION:
; APPLICANT: Yuejin Sun
; APPLICANT: Dilkes, Brian R.
; APPLICANT: Larkins, Brian A.
; APPLICANT: Lowe, Keith S.
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Dante, Ricardo
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/470,526
; CURRENT FILING DATE: 1999-12-22
; EARLIER APPLICATION NUMBER: 60/109,414
; EARLIER FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 9
; LENGTH: 525
; TYPE: PRT
; ORGANISM: zea mays
US-09-470-526-9

Query Match      88.6%; Score 31; DB 4; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GKGARR 6
Db      6 GKGARR 11

RESULT 10
US-09-252-991A-32826
; Sequence 32826, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32826
; LENGTH: 808
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32826

Query Match      88.6%; Score 31; DB 4; Length 808;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 1 GKGARR 6
| | | | |
Db 155 GKGARR 160

RESULT 11

US-09-627-851B-47
; Sequence 47, Application US/09627851B
; Patent No. 6545131
; GENERAL INFORMATION:
; APPLICANT: ISSACS, JOHN T.
; APPLICANT: DENMEADE, SAMUEL R.
; APPLICANT: CHRISTENSEN, S. BROGGER
; APPLICANT: LILJA, HANS
; TITLE OF INVENTION: TISSUE SPECIFIC PRODRUG
; FILE REFERENCE: 57111 (71699)
; CURRENT APPLICATION NUMBER: US/09/627, 851B
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/588, 822
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: 09/081, 707
; PRIOR FILING DATE: 1998-05-19
; PRIOR APPLICATION NUMBER: 60/047, 070
; PRIOR FILING DATE: 1997-05-19
; PRIOR APPLICATION NUMBER: 60/080, 046
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 47
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-627-851B-47

Query Match 85.7%; Score 30; DB 4; Length 7;
Best Local Similarity 85.7%; Pred. No. 4.1e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GKGARRL 7
| | | | |
Db 1 GKGARRL 7

RESULT 12

US-09-270-767-59934
; Sequence 59934, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270, 767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 59934
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-59934

Query Match 85.7%; Score 30; DB 4; Length 181;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GKGARRL 7
| | | | |
Db 65 GKGARRL 71

RESULT 13

US-09-252-991A-16910
; Sequence 16910, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252, 991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074, 788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094, 190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16910
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16910

Query Match 85.7%; Score 30; DB 4; Length 217;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GKGARRL 7
| | | | |
Db 14 GGSARRL 20

RESULT 14

US-09-252-991A-16592
; Sequence 16592, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252, 991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074, 788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094, 190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16592
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16592

Query Match 85.7%; Score 30; DB 4; Length 240;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GKGARRL 7
| | | | |
Db 234 GKGPRRL 240

RESULT 15

US-09-107-433-3048
; Sequence 3048, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
;
; INFORMATION FOR SEQ ID NO: 3048:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 251 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...251
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 3048:
US-09-107-433-3048

Query Match      85.7%; Score 30; DB 4; Length 251;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GKGARRL 7
      ||| |||
Db      133 GKGIRRL 139

RESULT 16
US-09-252-991A-21221
; Sequence 21221, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21221
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21221

Query Match      85.7%; Score 30; DB 4; Length 686;
```

```

      Best Local Similarity 85.7%; Pred. No. 4.6e+02;
      Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GKGARRL 7
      ||| |||
Db      275 GKGIRRL 281

RESULT 17
US-09-902-540-16744
; Sequence 16744, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16744
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-16744

Query Match      82.9%; Score 29; DB 4; Length 204;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GKARRL 7
      |||||
Db      195 GKARRL 200

RESULT 18
US-09-328-352-7258
; Sequence 7258, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7258
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7258

Query Match      82.9%; Score 29; DB 4; Length 217;
Best Local Similarity 71.4%; Pred. No. 2.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 GKGARRL 7
      ||: |||:
Db      23 GGOARRM 29

RESULT 19
US-09-252-991A-25021
; Sequence 25021, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
```



```
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25021
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25021

Query Match      82.9%; Score 29; DB 4; Length 233;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GKGARRL 7
Db      177 GGPARRL 183

RESULT 20
US-09-087-134-15
; Sequence 15, Application US/09087134
; Patent No. 6365711
; GENERAL INFORMATION:
; APPLICANT: Malcolm Whitman and Xin Chen
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
; TITLE OF INVENTION: TGF-BETA SUPERFAMILY SIGNALLING
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087,134
; FILING DATE: 27-MAY-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/047,991
; FILING DATE: 28-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Bieker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 00246/501002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Other
; LOCATION: 1...1
; OTHER INFORMATION: Human FAST-1 SID
US-09-087-134-15
```

```
Query Match      82.9%; Score 29; DB 3; Length 311;
Best Local Similarity 85.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GKGARRL 7
Db      152 GAKARRL 158

RESULT 21
US-09-949-016-10383
; Sequence 10383, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10383
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10383

Query Match      82.9%; Score 29; DB 4; Length 311;
Best Local Similarity 71.4%; Pred. No. 3.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 GKGARRL 7
Db      25 GGRSRRRL 31

RESULT 22
US-09-252-991A-21867
; Sequence 21867, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21867
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21867

Query Match      82.9%; Score 29; DB 4; Length 313;
Best Local Similarity 71.4%; Pred. No. 3.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 GKGARRL 7
Db      205 GGRARRV 211
```

```
RESULT 23
US-09-252-991A-21479
; Sequence 21479, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21479
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21479
```

```
Query Match      82.9%; Score 29; DB 4; Length 330;
Best Local Similarity 85.7%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 GKGARRL 7
      ||| |||
Db      117 GKGDRRL 123
```

```
RESULT 24
US-09-252-991A-28970
; Sequence 28970, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28970
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28970
```

```
Query Match      82.9%; Score 29; DB 4; Length 343;
Best Local Similarity 85.7%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 GKGARRL 7
      ||| |||
Db      315 GGAARRL 321
```

```
RESULT 25
US-09-252-991A-19721
; Sequence 19721, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
```

```
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19721
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19721
```

```
Query Match      82.9%; Score 29; DB 4; Length 372;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 GKGARRL 7
      ||| |||
Db      258 GKGARRL 264
```

```
RESULT 26
US-09-552-322-2
; Sequence 2, Application US/095523222
; Patent No. 6436642
; GENERAL INFORMATION:
; APPLICANT: Gould-Rothberg
; TITLE OF INVENTION: METHOD OF CLASSIFYING A THYROID CARCINOMA USING
; TITLE OF INVENTION: DIFFERENTIAL GENE EXPRESSION
; FILE REFERENCE: 15966-548
; CURRENT APPLICATION NUMBER: US/09/552,322
; CURRENT FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/130,123
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 60/193,203
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-552-322-2
```

```
Query Match      82.9%; Score 29; DB 4; Length 384;
Best Local Similarity 85.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 GKGARRL 7
      ||| |||
Db      24 GGPARRL 30
```

```
RESULT 27
US-09-252-991A-17341
; Sequence 17341, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17341
; LENGTH: 395
; TYPE: PRT
```

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17341

Query Match 82.9%; Score 29; DB 4; Length 395;
Best Local Similarity 85.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGAARL 7
|||
Db 33 GGPARRL 39

RESULT 28

US-09-902-540-11536
; Sequence 11536, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 11536
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-11536

Query Match 82.9%; Score 29; DB 4; Length 450;
Best Local Similarity 85.7%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGAARL 7
|||
Db 420 GGTARL 426

RESULT 29

US-09-252-991A-22080
; Sequence 22080, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22080
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22080

Query Match 82.9%; Score 29; DB 4; Length 454;
Best Local Similarity 71.4%; Pred. No. 5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAARL 7
|||
Db 38 GGRARRV 44

RESULT 30

US-09-252-991A-28585
; Sequence 28585, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28585
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28585

Query Match 82.9%; Score 29; DB 4; Length 458;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GKARL 7
|||
Db 397 GKARL 402

RESULT 31

US-09-561-763-2
; Sequence 2, Application US/09561763
; Patent No. 6664373
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J. et al.
; TITLE OF INVENTION: NOVEL POTASSIUM CHANNEL MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-074CP2
; CURRENT APPLICATION NUMBER: US/09/561,763
; CURRENT FILING DATE: 2000-04-29
; PRIOR APPLICATION NUMBER: 09/431,367
; PRIOR FILING DATE: 01-11-1999
; PRIOR APPLICATION NUMBER: US 09/259,951
; PRIOR FILING DATE: 01-03-1999
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-561-763-2

Query Match 82.9%; Score 29; DB 4; Length 499;
Best Local Similarity 71.4%; Pred. No. 5.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAARL 7
|||
Db 138 GGRAKRL 144

RESULT 32

US-09-431-367B-2
; Sequence 2, Application US/09431367B
; Patent No. 6670149
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: NOVEL POTASSIUM CHANNEL MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-074CP
; CURRENT APPLICATION NUMBER: US/09/431,367B
; CURRENT FILING DATE: 1999-11-01

```
; PRIOR APPLICATION NUMBER: 09/259,951
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-431-367B-2
```

```
Query Match      82.9%; Score 29; DB 4; Length 499;
Best Local Similarity 71.4%; Pred. No. 5.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGRARL 7
Db      138 GGRARL 144
```

RESULT 33

```
US-09-907-794A-315
; Sequence 315, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
```

```
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 315
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-907-794A-315
```

```
Query Match      82.9%; Score 29; DB 4; Length 509;
Best Local Similarity 85.7%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 GGRARL 7
Db      433 GGRARL 439
```

RESULT 34

```
US-09-905-125A-315
; Sequence 315, Application US/09905125A
; Patent No. 6664376
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,125A
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
```

```

; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 315
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-905-125A-315
```

```

Query Match      82.9%; Score 29; DB 4; Length 509;
Best Local Similarity 85.7%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 GGGKARL 7
Db      433 GGGKARL 439
```

```

RESULT 35
US-09-902-775A-315
; Sequence 315, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
```

```

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902, 775A
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143, 048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145, 698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146, 222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 315
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-902-775A-315
```

```

Query Match      82.9%; Score 29; DB 4; Length 509;
Best Local Similarity 85.7%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 GGGKARL 7
Db      433 GGGKARL 439
```

```

RESULT 36
US-09-906-700-315
; Sequence 315, Application US/09906700
; Patent No. 6723535
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
```



```
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,700
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 315
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-09-906-700-315

Query Match      82.9%; Score 29; DB 4; Length 509;
Best Local Similarity 85.7%; Pred. No. 5.5e+02;
Matches      6; Conservative      0; Mismatches      1; Indels      0; Gaps      0;
```

```
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: GNE.1618P2C12
; CURRENT APPLICATION NUMBER: US/09/903,603A
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 315
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-09-903-603A-315

Query Match      82.9%; Score 29; DB 4; Length 509;
Best Local Similarity 85.7%; Pred. No. 5.5e+02;
Matches      6; Conservative      0; Mismatches      1; Indels      0; Gaps      0;
```

OY 1 GGKARRL 7
| | | | |
Db 433 GGKARL 439

RESULT 38
US-09-904-920A-315
; Sequence 315, Application US/09904920A
; Patent No. 6806352
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904, 920A
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143, 048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145, 698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146, 222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 315
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-904-920A-315

Query Match 82.9%; Score 29; DB 4; Length 509;
Best Local Similarity 85.7%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGKARRL 7
| | | | |
Db 433 GGKARL 439

RESULT 39
US-09-909-064-315
; Sequence 315, Application US/09909064
; Patent No. 6818449
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909, 064
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143, 048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145, 698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146, 222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29

```
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 315
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-909-064-315

Query Match      82.9%; Score 29; DB 4; Length 509;
Best Local Similarity 85.7%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GGKARRL 7
Db      433 GGKARL 439

RESULT 40
US-09-905-381A-315
; Sequence 315, Application US/09905381A
; Patent No. 6818746
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,381A
; PRIOR APPLICATION NUMBER: 2001-07-13
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
```

```
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 315
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-905-381A-315

Query Match      82.9%; Score 29; DB 4; Length 509;
Best Local Similarity 85.7%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GGKARRL 7
Db      433 GGKARL 439

RESULT 41
US-09-906-618-315
; Sequence 315, Application US/09906618
; Patent No. 6828146
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
```

```
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,618
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 315
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-09-906-618-315

Query Match      82.9%; Score 29; DB 4; Length 509;
Best Local Similarity 85.7%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 GGAARL 7
        |||||
Db      433 GGAARL 439
```

```
RESULT 42
US-09-252-991A-28208
; Sequence 28208, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28208
; LENGTH: 514
; TYPE: PRT
```

```
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-28208

Query Match      82.9%; Score 29; DB 4; Length 514;
Best Local Similarity 85.7%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 GGAARL 7
        |||||
Db      14 GGAARL 20
```

```
RESULT 43
US-09-087-134-14
; Sequence 14, Application US/09087134
; Patent No. 6365711
; GENERAL INFORMATION:
; APPLICANT: Malcolm Whitman and Xin Chen
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
; TITLE OF INVENTION: TGF-BETA SUPERFAMILY SIGNALING
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087,134
; FILING DATE: 27-MAY-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/047,991
; FILING DATE: 28-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Bieker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 00246/501002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Other
; LOCATION: 1...1
; OTHER INFORMATION: Human FAST-1
; US-09-087-134-14

Query Match      82.9%; Score 29; DB 3; Length 544;
Best Local Similarity 85.7%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 GGAARL 7
        |||||
Db      385 GGAARL 391
```

```
RESULT 44
US-09-540-236-2071
; Sequence 2071, Application US/09540236
```

; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2071
; LENGTH: 579
; TYPE: PRT
; ORGANISM: M.catarhalis
US-09-540-236-2071

Query Match 82.9%; Score 29; DB 4; Length 579;
Best Local Similarity 71.4%; Pred. No. 6.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARRL 7
Db 502 GGOARRI 508

RESULT 45
US-09-252-991A-23627
; Sequence 23627, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23627
; LENGTH: 1413
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23627

Query Match 82.9%; Score 29; DB 4; Length 1413;
Best Local Similarity 85.7%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGKARRL 7
Db 1117 GGAARRL 1123

RESULT 46
US-08-504-265B-89
; Sequence 89, Application US/08504265B
; Patent No. 5837516
; GENERAL INFORMATION:
; APPLICANT: Marcus D. Ballinger and James A. Wells
; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING
; TITLE OF INVENTION: SUBSTRATES CONTAINING BASIC RESIDUES
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/504,265B
; FILING DATE: 19-Jul-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/398028
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0936P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8228
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-504-265B-89

Query Match 80.0%; Score 28; DB 2; Length 8;
Best Local Similarity 83.3%; Pred. No. 4.1e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGKARR 6
Db 3 GGAARR 8

RESULT 47
US-09-018-635-39
; Sequence 39, Application US/09018635
; Patent No. 6297356
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Broccoli, Dominique
; APPLICANT: Smogorzewska, Agata
; TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USE THEREOF
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KLAUBER & JACKSON
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/018,635
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: David A. Jackson
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-142 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: TRF2
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: mouse
US-09-018-635-39

Query Match 80.0%; Score 28; DB 3; Length 45;
Best Local Similarity 83.3%; Pred. No. 99;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARR 6
||:||||
Db 23 GGRARR 28

RESULT 48

US-09-912-962-39
Sequence 39, Application US/09912962
Patent No. 6586577

GENERAL INFORMATION:
APPLICANT: de Lange, Titia
Broccoli, Dominique
Smogorzewska, Agata
TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND
DIAGNOSTIC AND THERAPEUTIC USE THEREOF
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: KLAUBER & JACKSON
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/912,962
FILING DATE: 25-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/018,635
FILING DATE: 04-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: David A. Jackson
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-142 CIP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521

INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: TRF2
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: mouse
SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-09-912-962-39

Query Match 80.0%; Score 28; DB 4; Length 45;
Best Local Similarity 83.3%; Pred. No. 99;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARR 6
||:||||
Db 23 GGRARR 28

RESULT 49

US-09-902-540-12424
Sequence 12424, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 12424
LENGTH: 50
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-12424

Query Match 80.0%; Score 28; DB 4; Length 50;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKGARR 6
||:||||
Db 37 GKGARK 42

RESULT 50

US-09-270-767-40821
Sequence 40821, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 40821
LENGTH: 88
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-40821

Query Match 80.0%; Score 28; DB 4; Length 88;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GKGARRL 7
||:||||
Db 37 GKGARRM 43

Search completed: April 28, 2005, 06:26:58
Job time : 45 secs